

Hepatitis C Virus cell culture system

Publication number: EP1043399 (A2)

Publication date: 2000-10-11

Inventor(s): BARTENSCHLAGER RALF DR [DE] +

Applicant(s): BARTENSCHLAGER RALF DR [DE] +

Classification:

- international: C07K14/02; C12N15/09; C12N15/86; C12N5/08; C12N5/10; C12N7/00; C12N7/01; A61K39/00; A61K48/00; C12R1/91; C07K14/005; C12N15/09; C12N15/86; C12N5/08; C12N5/10; C12N7/00; C12N7/01; A61K39/00; A61K48/00; (IPC1-7); A61K48/00; A61K49/00; C07K14/18; C12N15/86; C12N5/10; C12N7/01; C12N7/04

- European: C07K14/02; C12N15/86; C12N7/00

Application number: EP2000010529 20000323

Priority number(s): DE19991015178 19990403

Also published as:

 EP1043399 (A3)
 EP1043399 (B1)
 EP1043399 (B9)
 DE199151178 (A1)
 US6630343 (B1)

[more >>](#)

Cited documents:

 WO9839031 (A1)
 WO9904008 (A2)
 US5851758 (A)
 US5874565 (A)
 WO9967394 (A1)

Abstract of EP 1043399 (A2)

Hepatitis C virus (HCV) cell culture system comprising human hepatoma cells that contain an integrated HCV-RNA construct (I) new. (I) contains the HCV-specific RNA segments 5'-NTR (non-translated region), NS (non-structural)3, NS4A, NS4B, NS5A, NS5B and 3'-NTR, and a selectable (marker) gene (II). An independent claim is also included for (I) containing the novel HCV-specific RNA segments.

 Data supplied from the **espacenet** database — Worldwide



(11) **EP 1 043 399 A2**

(12) **EUROPÄISCHE PATENTANMELDUNG**

(43) Veröffentlichungstag:
11.10.2000 Patentblatt 2000/41

(51) Int. Cl.⁷: **C12N 15/86**, C12N 7/01,
C12N 7/04, C12N 5/10,
C07K 14/18, A61K 49/00,
A61K 48/00

(21) Anmeldenummer: 00105929.4

(22) Anmeldetag: 23.03.2000

(84) Benannte Vertragsstaaten:
AT BE CH CY DE DK ES FI FR GB GR IE IT LI LU
MC NL PT SE
Benannte Erstreckungsstaaten:
AL LT LV MK RO SI

(72) Erfinder:
Bartenschlager, Ralf, Dr.
55239 Gau-Odernheim (DE)

(30) Priorität: 03.04.1999 DE 19915178

(74) Vertreter: **Rudolph, Ulrike, Dr.**
Patentanwältin
In der Schanz 10
69198 Schriesheim (DE)

(71) Anmelder:
Bartenschlager, Ralf, Dr.
55239 Gau-Odernheim (DE)

(54) **Hepatitis C Virus Zellkultursystem**

(57) Das erfindungsgemäße Hepatitis C Virus (HCV) Zellkultursystem besteht aus humanen Hepato-mazellen, die mit einem HCV-RNA-Konstrukt transfiziert sind, das die HCV-spezifischen RNA-Abschnitte 5' NTR, NS3, NS4A, NS4B, NS5A, NS5B und 3' NTR und zudem wenigstens ein selektierbares Markergen (Selektionsgen) umfaßt.

EP 1 043 399 A2

Beschreibung

[0001] Die Erfindung betrifft ein Hepatitis C Virus (HCV) Zellkultursystem, das im wesentlichen eukaryontische Zellen umfaßt, die eingeschleustes HCV-spezifisches Genmaterial enthalten, d.h. die mit HCV-spezifischem Genmaterial transfiguriert sind.

[0002] Das Hepatitis C Virus (HCV) ist eine der Hauptursachen chronischer und sporadischer Leberkrankungen weltweit. Die meisten HCV-Infektionen verlaufen ohne erkennbare klinische Symptome, allerdings werden 80-90% der Infizierten dauerhafte Virusträger und bei 50% dieser dauerhaften Virusträger kommt es zu einer chronischen Leberentzündung mit unterschiedlichen Schweregraden. Ca. 20% der chronisch Infizierten entwickeln im Laufe von 10 bis 20 Jahren eine Leberzirrhose, auf deren Basis sich ein primäres Leberzellkarzinom entwickeln kann. Die chronische Hepatitis C ist heute die Hauptindikation für eine Lebertransplantation. Eine Kausaltherapie gibt es bisher noch nicht. Die einzige derzeit verfügbare Therapie ist die hochdosierte Verabreichung von Interferon-Alpha oder eine Kombination aus Interferon-Alpha und dem Purin-Nukleosidanalogon Ribavirin. Allerdings sprechen nur ca. 60 % aller Behandelten auf diese Therapie an und bei diesen kommt es in mehr als der Hälfte aller Fälle nach dem Absetzen der Behandlung zu einer erneuten Virämie.

Aufgrund der hohen Prävalenz, gerade auch in den Industrieländern, den schwerwiegenden Folgen chronischer Infektionen und dem Fehlen einer Kausaltherapie ist die Entwicklung einer HCV-spezifischen Chemotherapie ein wesentliches Ziel der pharmazeutischen Forschung und Entwicklung. Hauptproblem hierbei ist bisher das Fehlen eines geeigneten Zellkultursystems, das ein Studium der Virus-Replikation und der Pathogenese in eukaryontischen Zellen ermöglicht.

[0003] Aufgrund der geringen Virusmengen im Blut bzw. Gewebe, dem Fehlen geeigneter Zellkultursysteme oder Tiermodelle (bis heute ist der Schimpanse das einzige mögliche Versuchstier) sowie dem Fehlen effizienter Systeme zur Produktion virus-ähnlicher Partikel, konnte die molekulare Zusammensetzung des HCV-Partikels bis heute noch nicht eingehend untersucht bzw. aufgeklärt werden. Die derzeit vorliegenden Ergebnisse lassen sich wie folgt zusammenfassen: Das HCV ist ein unhülltes Plusstrang RNA Virus mit einem Partikeldurchmesser von 50-60 nm und einer mittleren Dichte von 1,03-1,1g/ml. Es wurde erstmals 1989 molekular kloniert und charakterisiert (Choo et al., 1989; Science, 244, 359-362). Die HCV-RNA hat eine Länge von ca. 9.6 kb (= 9600 Nukleotide), eine positive Polarität und besitzt ein einziges offenes Leseraster (ORF = open reading frame), das ein lineares Polyprotein von ca 3010 Aminosäuren kodiert (siehe Rice 1996, in Virology, B. N. Fields, D. M. Howley, Eds. (Lippincott-Raven, Philadelphia, PA, 1996), vol. 1, pp.931-960; Clarke 1997, J. Gen. Virol. 78, 2397; und Bartschlagler 1997, Intervirology 40, 378 und vgl. Fig. 1 A). Bei der Virusreplikation wird das Polyprotein durch zelluläre und virale Proteasen in die reifen und funktionell aktiven Proteine gespalten.

Innerhalb des Polyproteins sind die Proteine wie folgt angeordnet (vom Amino- zum Carboxyterminus): Core-E1-E2-p7-NS2-NS3-NS4A-NS4B-NS5A-NS5B. Das Core-Protein ist die Hauptkomponente des Nukleokapsids. Die Glykoproteine E1 und E2 sind Transmembranproteine und Hauptkomponenten der Virushülle. Sie spielen wahrscheinlich bei der Anheftung des Virus an die Wirtszelle eine wesentliche Rolle. Diese drei Proteine Core, E1 und E2 haben den Viruspartikel auf und werden deshalb als Strukturproteine bezeichnet. Die Funktion des Proteins p7 ist noch unklar. Das Protein NS2 ist wahrscheinlich die katalytische Domäne der NS2-3 Protease, die für die Prozessierung zwischen den Proteinen NS2 und NS3 verantwortlich ist. Das Protein NS3 hat zwei Funktionen, nämlich in der aminoterminalen Domäne eine Proteaseaktivität, die für die Polyproteinprozessierung essentiell ist, und in der carboxyterminalen Domäne eine NTPase/Helikase-Funktion, die wahrscheinlich bei der Replikation der viralen RNA eine Rolle spielt. Das Protein NS4A ist ein Kofaktor der NS3-Protease. Die Funktion des Proteins NS4B ist unbekannt.

[0004] Das offene Leseraster ist an seinem 5' Ende von einer ca 340 Nukleotide langen nicht-translatierten Region (NTR = non-translated region) flankiert, die als interne Ribosomenansatzstelle (IRES = internal ribosome entry site) fungiert, und an seinem 3' Ende von einer ca. 230 Nukleotide langen NTR, die höchstwahrscheinlich für die Genomreplikation von Bedeutung ist. Eine solche 3' NTR ist Gegenstand der Patentanmeldung PCT/US 96/14033. Die Strukturproteine in dem amino-terminalen Viertel des Polyproteins werden von der Signalpeptidase der Wirtszelle gespalten. Die Nicht-Strukturproteine (NS) 2 bis (NS) 5B werden von zwei viralen Enzymen prozessiert, nämlich von der NS2-3 und der NS3/4A Proteinase. Die NS3/4A Proteinase wird für alle Spaltungen jenseits des Carboxyterminus von NS3 benötigt. Die Rolle von NS4B ist nicht bekannt. NS5A, ein hoch phosphoryliertes Protein, scheint für die Interferon Resistenz verschiedener HCV-Genotypen verantwortlich zu sein (vgl. Enomoto et al. 1995, J. Clin. Invest. 96, 224; Enomoto et al. 1996, N. Engl. J. Med. 334, 77; Gale Jr. et al. 1997, Virology 230, 217; Kaneko et al. 1994, Biochem. Biophys. Res. Commun. 205, 320; Reed et al., 1997, J. Virol. 71, 7187) und NS5B wurde als die RNA-abhängige RNA Polymerase identifiziert.

[0005] Anhand dieser Erkenntnisse wurden erste Diagnosesysteme entwickelt, die entweder auf dem Nachweis von HCV-spezifischen Antikörpern in Patientenserum oder auf dem Nachweis von HCV-spezifischer RNA mittels RT-PCR (= Reverse Transcription Polymerase Chain Reaction) beruhen, und die mittlerweile routine- und/oder vorschriftsmäßig bei allen Blutkonserven angewendet werden (müssen).

[0006] Seit der Erstbeschreibung des Genoms 1989 wurden mit Hilfe der PCR-Methode zahlreiche Teil- und Komplettssequenzen des HCV kloniert und charakterisiert. Ein Vergleich dieser Sequenzen zeigt eine hohe Variabilität des viralen Genoms, insbesondere im Bereich des NS5B-Gens, was letztendlich zu einer Einteilung in 6 Genotypen geführt hat, die selbst nochmals in Subtypen a, b, und c untergliedert sind. Die genomische Varianz ist nicht gleichmäßig über das Genom verteilt. So sind die 5'NTR und Teile der 3'NTR hoch konserviert, während bestimmte kodierende Sequenzen z.T. sehr stark variieren, vor allem die Hüllproteine E1 und E2.

[0007] Die klonierten und charakterisierten Teil- und Komplettssequenzen des HCV-Genoms wurden außerdem hinsichtlich geeigneter Angriffsziele für ein prospektives antivirales Therapeutikum untersucht. Dabei wurden drei virale Enzyme gefunden, die sich als solches Angriffsziel anbieten. Diese sind (1) der NS3/4A Proteasekomplex, (2) die NS3 Helikase und (3) die NS5B RNA-abhängige RNA Polymerase. Der NS3/4A Proteasekomplex und die NS3 Helikase konnten bereits kristallisiert und hinsichtlich ihrer dreidimensionalen Struktur aufgeklärt werden (Kim et al., 1996, Cell, 87,343; Yem et al., 1998, Protein Science, 7, 837; Love et al., 1996, Cell, 87, 311; Kim et al., 1998, Structure, 6, 89; Yao et al., 1997, Nature Structural Biology, 4, 463; Cho et al., 1998, J. Biol. Chem., 273, 15045); bei der NS5B RNA-abhängigen RNA Polymerase ist dies bis heute noch nicht gelungen.

Obwohl mit diesen Enzymen bedeutsame Angriffsziele für eine Therapieentwicklung der chronischen HCV-Infektion definiert sind, und obwohl sowohl mit Hilfe von 'rational drug design' als auch mit Hilfe von 'high throughput screens' weltweit intensiv nach geeigneten Inhibitoren gesucht wird, leidet die Therapieentwicklung an einem großen Defizit, nämlich dem Fehlen von Zellkultursystemen oder einfachen Tiermodellen, die es erlauben, HCV-RNA oder HCV-Antigene direkt, zuverlässig und mit einfachen laborüblichen Methoden nachzuweisen. Das Fehlen solcher Zellkultursysteme ist auch der Hauptgrund dafür, daß das Verständnis der HCV-Replikation bis heute noch sehr lückenhaft und in weiten Teilen nur hypothetisch ist.

[0008] Obwohl nach Meinung der Fachwelt eine enge evolutionäre Beziehung zwischen HCV und den Flavi- und Pestiviren besteht und für diese autonom replizierende RNAs beschrieben sind, die in verschiedenen Zelllinien ohne weiteres zur Replikation gebracht werden können und dabei relativ hohe Ausbeuten zeigen (siehe Khromykh et al., 1997, J. Virol. 71, 1497; Behrens et al., 1998, J. Virol. 72, 2364; Moser et al., 1998, J. Virol. 72, 5318), waren ähnliche Versuche mit HCV bisher nicht erfolgreich.

[0009] Zwar ist aus verschiedenen Publikationen bekannt, daß Zelllinien oder primäre Zellkulturen mit HCV-haltigem, hochtrigem Patientenserum infiziert werden können, (Lanford et al. 1994, Virology 202, 606; Shimizu et al. 1993, *Proceedings of the National Academy of Sciences, USA*, 90, 6037-6041; Mizutani et al. 1996, *Journal of Virology*, 70, 7219-7223; M. Ikeda et al. 1998, *Virus Res.* 56, 157; Fournier et al. 1998, *J. Gen. Virol.* 79, 2376 und darin zitierte Literaturstellen, Ito et al. 1996, *Journal of General Virology*, 77, 1043-1054), diese virusinfizierten Zelllinien oder Zellkulturen erlauben jedoch nicht den direkten Nachweis von HCV-RNA oder HCV-Antigenen. Die virale RNA in diesen Zellen ist weder in einem Northern-Blot (einem Standardverfahren zum quantitativen Nachweis von RNA) noch sind die viralen Proteine in einem Western-Blot oder mittels Immunpräzipitation detektierbar. Nur mit sehr aufwendigen und indirekten Methoden ist es überhaupt gelungen, eine HCV-Replikation nachzuweisen. Diese nachteiligen Umstände zeigen klar, daß die Replikation in diesen bekannten virusinfizierten Zelllinien oder Zellkulturen absolut unzureichend ist.

[0010] Desweiteren ist aus den Publikationen von Yoo et al. (1995, *Journal of Virology*, 69, 32-38) und von Dash et al., (1997, *American Journal of Pathology*, 151, 363-373) bekannt, daß Hepatomazelllinien mit synthetischer HCV-RNA, die mittels in vitro Transkription von kloniertem HCV-Genom gewonnen wurde, transfiziert werden können. In beiden Publikationen gingen die Autoren von dem Grundgedanken aus, daß das virale HCV-Genom eine Plusstrang-RNA ist, die nach dem Einschleusen in die Zelle direkt als mRNA fungiert, an die sich Ribosomen anheften und im Zuge von Translationsprozessen Virusproteine bilden, aus denen sich letztendlich neue HCV-Partikel bilden (können). Diese Virusreplikation, d.h. diese neu gebildeten HCV-Viren bzw. deren RNA wurde mittels RT-PCR nachgewiesen. Die publizierten Ergebnisse der durchgeführten RT-PCR sprechen jedoch dafür, daß die Effizienz der HCV-Replikation in den beschriebenen HCV-transfizierten Hepatomazellen nur sehr gering ist und jedenfalls nicht ausreicht, um Schwankungen in der Replikationsrate nach gezielter Einwirkung mit prospektiven antiviralen Therapeutika auch nur qualitativ, geschweige denn quantitativ zu messen. Außerdem ist im Stand der Technik mittlerweile bekannt (Yanagi et al., *Proc. Natl. Acad. Sci. USA*, 96, 2291-95, 1999), daß die hochkonservierte 3' NTR essentiell ist für die Virusreplikation, was in klarem Widerspruch zu den Behauptungen von Yoo et al. und Dash et al. steht, die für ihre Versuche in Unkenntnis des authentischen 3' Endes des HCV-Genoms ausschließlich HCV-Genome mit verkürzten 3' NTRs verwendet haben.

[0011] Aufgabe der vorliegenden Erfindung ist die Bereitstellung eines HCV-Zellkultursystems, bei dem die virale RNA in den transfizierten Zellen autonom und mit so hoher Effizienz repliziert, daß Schwankungen in der Replikationsrate nach gezielter Einwirkung mit virus- und insbesondere HCV-spezifischen prospektiven antiviralen Therapeutika qualitativ und quantitativ und mit Hilfe gängiger, laborüblicher Meßverfahren gemessen werden können.

[0012] Eine Lösung dieser Aufgabe besteht in der Bereitstellung eines Zellkultursystems der eingangs genannten Art, bei dem die eukaryontischen Zellen humane Zellen, insbesondere Hepatomazellen sind, die vorzugsweise von einer handelsüblichen Hepatomazelllinie abstammen, aber auch aus einer entsprechenden Primärzellkultur gewonnen sein können, und bei dem das eingeschleuste HCV-spezifische Genmaterial ein HCV-RNA-Konstrukt ist, das im

wesentlichen die HCV-spezifischen RNA-Abschnitte 5' NTR, NS3, NS4A, NS4B, NS5A, NS5B und 3' NTR, vorzugsweise in der genannten Reihenfolge, und zudem wenigstens ein selektierbares Markergen (Selektionsgen) umfaßt. "NTR" steht hier und im folgenden für "nicht-translatierte Region" und ist dem einschlägigen Fachmann als Begriff bzw. Abkürzung bekannt und geläufig. Der Begriff "HCV-RNA-Konstrukt" umfaßt hier und im folgenden sowohl Konstrukte, die das komplette HCV-Genom enthalten, als auch solche, die lediglich einen Teil davon, d.h. ein HCV-Subgenom enthalten.

Eine bevorzugte Variante des erfindungsgemäßen Zellkultursystems, die sich in der Praxis sehr gut bewährt hat, ist unter der Nummer DSM ACC2394 (Laborbezeichnung HuBI 9-13) bei der DSMZ, Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH in Braunschweig, Deutschland, hinterlegt.

- [0013] Mit dem erfindungsgemäßen Zellkultursystem wird erstmals ein in-vitro-System bereit gestellt, in dem HCV-RNA intrazellulär, autonom und in ausreichend großen Mengen repliziert und exprimiert wird, so daß eine quantitative Bestimmung sowohl der HCV-RNA-Mengen als auch der HCV-spezifischen Proteine mit konventionellen und zuverlässig genauen biochemischen Meßmethoden durchgeführt werden kann. Das heißt: es steht erstmals ein annähernd authentisches zellgestütztes ("cell-based") HCV-Replikationssystem zur Verfügung, das für die Entwicklung und Erprobung von antiviralen Pharmazeutika dringend benötigt wird. Dieses Testsystem bietet nun die Möglichkeit, potentielle Angriffspunkte für eine wirksame HCV-spezifische Therapie zu identifizieren und HCV-spezifische Chemotherapeutika zu entwickeln und zu evaluieren.

- [0014] Die Erfindung basiert auf der überraschenden Erkenntnis, daß eine effiziente Replikation der HCV-RNA nur dann in Zellen stattfindet, wenn diese mit einem HCV-RNA-Konstrukt transfiziert wurden, das mindestens die 5' und die 3' nicht-translatierten Regionen (NTR) und die Nichtstrukturproteine (NS) 3 bis 5B umfaßt und zusätzlich ein selektierbares Markergen (Selektionsgen) aufweist. Offensichtlich sind die Strukturgene für den Ablauf der Replikation ohne wesentliche Bedeutung, während andererseits eine effiziente Replikation der HCV-RNA anscheinend nur dann stattfindet, wenn die transfizierten Zellen einem permanenten Selektionsdruck unterzogen werden, der durch das mit der HCV-RNA verbundene selektierbare Markergen (Selektionsgen) vermittelt wird. Das Markergen (Selektionsgen) scheint somit einerseits die Selektion derjenigen Zellen zu provozieren, in denen die HCV-RNA produktiv repliziert, und andererseits scheint es die Effizienz der RNA-Replikation wesentlich zu steigern.

[0015] Gegenstand der Erfindung ist auch ein zellfreies HCV-RNA-Konstrukt, das sich dadurch auszeichnet, daß es die HCV-spezifischen RNA-Abschnitte 5' NTR, NS3, NS4A, NS4B, NS5A, NS5B und 3' NTR, vorzugsweise in der genannten Reihenfolge, und zudem ein selektierbares Markergen (Selektionsgen) umfaßt.

- [0016] Der Begriff 5' NTR bzw. NS3 bzw. NS4A bzw. NS4B bzw. NS5A bzw. NS5B bzw. 3' NTR umfaßt im vorliegenden Zusammenhang jede Nukleotidsequenz, die im Stand der Technik als Nukleotidsequenz für den jeweils betreffenden funktionellen Abschnitt des HCV-Genoms beschrieben ist.

- [0017] Die Bereitstellung eines solchen HCV-RNA-Konstrukts ermöglicht erstmals eine detaillierte Analyse der HCV - Replikation, - Pathogenese und - Evolution in Zellkulturen. Die HCV-spezifische virale RNA kann - als vollständiges Genom oder als Subgenom - gezielt in beliebigen Mengen erzeugt werden, und es besteht die Möglichkeit, das RNA-Konstrukt zu manipulieren und damit die HCV-Funktionen auf genetischer Ebene zu untersuchen und aufzuklären.

- [0018] Da alle zur Zeit als Hauptangriffsziel für eine Therapie untersuchten HCV-Enzyme, nämlich die NS3/4A Protease, die NS3 Helikase und die NS5B Polymerase, in dem erfindungsgemäßen HCV-RNA-Konstrukt enthalten sind, kann es für alle entsprechenden Untersuchungen benutzt werden.

- [0019] Eine Ausführungsform des HCV-RNA-Konstrukts, die sich in der praktischen Anwendung sehr gut bewährt hat, zeichnet sich dadurch aus, daß sie die Nukleotidsequenz gemäß Sequenzprotokoll SEQ ID NO:1 umfaßt. Weitere Ausführungsvarianten mit vergleichbar guten Eigenschaften für den Einsatz in der Praxis sind dadurch gekennzeichnet, daß sie eine Nukleotidsequenz entweder gemäß Sequenzprotokoll SEQ ID NO:2 oder SEQ ID NO:3 oder SEQ ID NO:4 oder SEQ ID NO:5 oder SEQ ID NO:6 oder SEQ ID NO:7 oder SEQ ID NO:8 oder SEQ ID NO:9 oder SEQ ID NO:10 oder SEQ ID NO:11 umfassen.

- [0020] Es besteht die Möglichkeit, das erfindungsgemäße HCV-Subgenom-Konstrukt mit einer 3' NTR zu versehen, die eine im Stand der Technik hierfür bisher unbekannte Nukleotidsequenz aufweist, nämlich eine Nukleotidsequenz, die aus der Gruppe der nachfolgend aufgelisteten Nukleotidsequenzen (a) bis (i) ausgewählt ist:

- 5 (a) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTTTTTTTA
GCTTTTTTTTTTCTTTTTTTTGAGAGAGAGAGTCTCACTCTGTGCCC
AGACTGGAGT
- 10 (b) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTT
TTTTTAGTCT TTTTTTTTCT TTTTTTTTGA GAGAGAGAGT CTCACTCTGT
TGCCAGACT GGAGC
- 15 (c) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTT
TTTAATCTTT TTTTTTTTCT TTTTTTTTGA GAGAGAGAGT CTCACTCTGT
TGCCAGACT GCAGC
- 20 (d) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTT
TTTTTAGTC TTTTTTTTTT TCTTTTTTTT TGAGAGAGAG AGTCTCACT
TGTTGCCAG ACTGGAGT
- 25 (e) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTT
TTTTTAGTCT TTTTTTTTTT TCTTTTTTTT TGAGAGAGAG AGTCTCACT
TGTTGCCAG ACTGGAGT
- 30 (f) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTT
TTTTTAGTCT TTTTTTTTTT TCTTTTTTTT TTGAGAGAGA GAGTCTCACT
CTGTGCCCA GACTGGAGT
- 35 (g) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTT
TTTTTAGTCT TTTTTTTTTT CTTTTTTTTT GAGAGAGAGA
GTCTCACTCT GTTGCCAGA CTGGAGT
- 40 (h) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTT
TTTTTTAAT CTTTTTTTTT TTTTCTTT TTTTGAGAGA
GAGAGTCTCA CTCTGTTGCC CAGACTGGAG T
- 45 (i) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTT
TTTTTTAATC TTTTTTTTTT TTTTCTTTT TTTTGAGAG
AGAGAGTCTC ACTCTGTG CAGACTGGA GT

50 Das in den erfindungsgemäßen HCV-RNA-Konstrukten enthaltene selektierbare Markergen (Selektionsgen) ist vorzugsweise ein Resistenzgen, insbesondere eine Antibiotikumresistenzgen.

Das hat den Vorteil, daß die mit diesem Konstrukt transfizierten Zellen leicht von den nicht transfizierten Zellen selektiert werden können, indem dem Zellkulturmedium z.B. im Fall eines Antibiotikumresistenzgens das betreffende Antibiotikum zugegeben wird. Unter 'Antibiotikum' wird im vorliegenden Zusammenhang jede Substanz verstanden, die die nicht-transfizierten Wirtszellen oder die Zellen, in denen die HCV-RNA nur mit geringer Effizienz repliziert, am Leben oder Wachstum hindert, insbesondere Zellgifte wie z.B. Puromycin, Hygromycin, Zeocin, Bleomycin oder Blastidin.

55 [0021] Ein bevorzugtes selektierbares Markergen (Selektionsgen) bzw. Resistenzgen, das sich in der Praxis sehr gut bewährt hat, ist das Neomycinphosphotransferasegen.

[0022] Eine Alternative zu den Antibiotikumresistenzgenen ist z.B. das Thymidin-Kinase-Gen, mit dem eine HAT-

Selektion durchgeführt werden kann.

[0023] Die Position des selektierbaren Markergens (Selektionsgens), bzw. des bevorzugten Resistenzgens bzw. des besonders bevorzugten Antibiotikumresistenzgens in dem HCV-RNA-Konstrukt liegt vorzugsweise hinter der HCV 5' NTR, d.h. strängabwärts der 5' NTR bzw. strängaufwärts des HCV-Leserasters. Denkbar ist aber auch eine Insertion im Bereich der 3' NTR oder an anderer Stelle des HCV-Genoms oder -Subgenoms, z.B. innerhalb des Polyproteins.

[0024] Bei einer alternativen Ausführungsform des erfindungsgemäßen HCV-RNA-Konstrukts ist das selektierbare Markergen (Selektionsgen), insbesondere ein Antibiotikumresistenzgen, über ein Ribozym bzw. eine Erkennungsstelle für ein Ribozym mit der HCV-RNA bzw. der HCV-Genom- oder -Subgenomsequenz verbunden.

[0025] Damit geht der Vorteil einher, daß nach erfolgter Selektion derjenigen Zellen, in denen die HCV-RNA produktiv repliziert, in den daraus gewonnenen Zellklonen das Resistenzgen durch ribozymvermittelte Spaltung von der HCV-Subgenomsequenz abgetrennt werden kann, nämlich durch Aktivierung des einklonierten Ribozyms oder, im Fall eines Konstrukts mit einer Erkennungsstelle für ein Ribozym, durch Einschleusen des Ribozyms in die Zellen (z.B. mittels Transfektion eines Ribozymkonstrukts oder Infektion mit einem viralen Expressionsvektor, in den das entsprechende Ribozym eingesetzt wurde). Auf diese Weise wird ein authentisches HCV-Genom-Konstrukt ohne Resistenzgen erhalten, das zur Bildung authentischer infektiöser Viruspartikel befähigt ist.

[0026] Eine weitere bevorzugte Ausführungsform des erfindungsgemäßen HCV-RNA-Konstrukts zeichnet sich dadurch aus, daß das Konstrukt wenigstens ein integriertes Reportergen aufweist.

[0027] Unter Reportergen wird im folgenden jedes Gen verstanden, dessen Anwesenheit sich nach Überführung in einen Zielorganismus leicht und im allgemeinen mit einfachen biochemischen oder auch histochemischen Methoden nachweisen läßt, d.h. das für ein Protein kodiert, welches auch in geringen Mengen einfach und zuverlässig mit den laborüblichen Meßmethoden nachgewiesen und quantifiziert werden kann.

[0028] Diese Variante des HCV-RNA-Konstrukts hat den Vorteil, daß der Umfang der Replikation dieses Konstrukts anhand des Reportergenprodukts einfach und schnell mit laborüblichen Methoden gemessen werden kann.

[0029] Das Reportergen ist vorzugsweise ein Gen aus der Gruppe der Luziferasene, dem CAT-Gen (Chloramphenicol-Acetyl-Transferase-Gen), dem lacZ-Gen (beta-Galaktosidasegen), dem GFP-Gen (green-fluorescence-protein-Gen), dem GUS-Gen (Glukuronidasegen) oder dem SEAP-Gen (Seszernte-Alkalische-Phosphatase-Gen). Diese Reportergene bzw. deren Produkte, nämlich die entsprechenden Reporterproteine, können z.B. mittels Fluoreszenz, Chemilumineszenz, colorimetrisch oder mit Hilfe immunologischer Methoden (z.B. ELISA) bestimmt werden.

[0030] Als Reportergen kommt aber auch ein Surrogatmarkergen in Betracht. Darunter sind in diesem Zusammenhang solche Gene zu verstehen, die für zelluläre Proteine, Nukleinsäuren oder — allgemein — für solche Funktionen kodieren, die einer von der Virusreplikation abhängigen Variation unterliegen, und die infolgedessen in denjenigen Zellen, in denen sich das HCV bzw. das HCV-RNA-Konstrukt vermehrt, entweder reprimiert oder aktiviert werden. Das heißt: die Reduktion bzw. Aktivierung dieser Funktion ist ein Ersatzmarker für die Virusreplikation bzw. die Replikation des HCV-RNA-Konstrukts.

[0031] Die Positionen von Reportergen und selektierbarem Markergen (Selektionsgen) können so gewählt sein, daß ein aus den beiden Genprodukten gebildetes Fusionsprotein exprimiert wird. Hierbei besteht die vorteilhafte Möglichkeit, daß diese beiden Gene so in dem HCV-RNA-Konstrukt angeordnet sind, daß ihre beiden exprimierten Proteine zunächst über eine Schnittstelle für eine Protease (z.B. Ubiquitin) oder über ein selbstspaltendes Peptid (z.B. das 2A-Protein der Picornaviren) fusioniert sind und erst später proteolytisch wieder getrennt werden.

[0032] Ebenso gut können diese beiden Positionen aber auch derart getrennt voneinander liegen, daß beide Genprodukte separat exprimiert werden. (z.B. in der Reihenfolge: Marker- bzw. Resistenzgen — interne Ribosomenbindungsstelle — Reportergen).

Im Fall des Reportergens hat sich eine Ausführungsvariante besonders bewährt, bei der das Reportergen in das offene Leseraster des HCV-Genoms oder -Subgenoms einkloniert ist, und zwar derart, daß es erst nach einer proteolytischen Prozessierung in eine aktive Form überführt wird.

[0032] Das erfindungsgemäße Zellkultursystem in allen seinen Variationen kann für vielfältige Zwecke eingesetzt werden. Diese umfassen:

- Das Auffinden antiviral wirksamer Substanzen. Dies können beispielsweise sein: organische Verbindungen, die unmittelbar oder mittelbar in die Virusvermehrung eingreifen (z.B. Inhibitoren der viralen Proteasen, der NS3-Helikase, der NS5B RNA-abhängigen RNA Polymerase), antisense Oligonukleotide, die an eine beliebige Zielsequenz innerhalb des HCV-RNA-Konstrukts (z.B. die 5' NTR) hybridisieren und unmittelbar oder mittelbar zu einer Beeinflussung der Virusvermehrung führen z.B. auf Grund einer Reduktion der Translation des HCV-Polyproteins oder Ribozyme, die eine beliebige HCV-RNA-Sequenz spalten und damit die Virusreplikation beeinträchtigen.
- Die Evaluierung jeglicher Art antiviral wirksamer Substanzen in Zellkultur. Solche Substanzen können beispielsweise mittels 'rational drug design' oder 'high-throughput screening' am isolierten gereinigten Enzym gefunden werden. Unter Evaluierung sind vor allem die Bestimmung der inhibitorischen Eigenschaften der entsprechenden Substanz sowie deren Wirkungsmechanismus zu verstehen.

- Die Identifikation neuer Angriffsziele, viralen oder zellulären Ursprungs, für eine HCV-spezifische antivirale Therapie. Ist beispielsweise ein zelluläres Protein essentiell für die Virusreplikation, kann mittels Hemmung dieses zellulären Proteins die Virusreplikation ebenfalls beeinflusst werden. Das Auffinden solcher auxiliären Faktoren ist mit dem erfindungsgemäßen System ebenfalls möglich.
- 5 • Der Einsatz für die Resistenzbestimmung. Es ist anzunehmen, daß auf Grund der hohen Mutationsrate des HCV-Genoms Therapieresistenzen auftreten. Solche Resistenzen, die gerade bei der klinischen Zulassung einer Substanz von großer Bedeutung sind, lassen sich mit dem erfindungsgemäßen Zellsystem ermitteln. Zelllinien in denen sich das HCV-RNA-Konstrukt bzw. das HCV-Genom oder - Subgenom repliziert, werden mit steigenden Konzentrationen der entsprechenden Substanz inkubiert und die Replikation der viralen RNA wird entweder
- 10 anhand eines eingebrachten Reporters oder durch qualitative oder quantitative Bestimmung der viralen Nukleinsäuren oder Proteine bestimmt. Resistenz ist dann gegeben, wenn bei normaler Wirkstoffkonzentration keine Hemmung der Replikation zu beobachten ist. Durch Reklonierung der HCV-RNA (z.B. mittels RT-PCR) und Sequenzanalyse können die für Therapieresistenz verantwortlichen Nukleotid- bzw. Aminosäureaustausche ermittelt werden. Durch Einklonieren der/des entsprechenden Austausch(es) in das Ursprungs-konstrukt kann deren Kausalität für die Therapieresistenz bewiesen werden.
- Die Produktion von authentischen Virusproteinen (Antigene) für die Entwicklung und/oder Evaluierung von Diagnostika. Das erfindungsgemäße Zellsystem erlaubt auch die Expression von HCV-Antigenen in Zellkulturen. Diese Antigene können prinzipiell auch für den Aufbau diagnostischer Nachweisverfahren eingesetzt werden.
- Die Produktion von HCV Viren und virus-ähnlichen Partikeln insbesondere zur Entwicklung oder Herstellung von
- 20 Therapeutika und Impfstoffen sowie für diagnostische Zwecke. Insbesondere zellkultur-adaptierte vollständige HCV-Genome, die mit dem erfindungsgemäßen Zellsystem hergestellt werden können, sind in der Lage, mit hoher Effizienz in Zellkulturen zu replizieren. Diese Genome besitzen alle Funktionen des HCV und sind deshalb in der Lage infektiöse Viren zu produzieren.
- 25 **[0033]** Das erfindungsgemäße HCV-RNA-Konstrukt für sich genommen kann in allen seinen Variationen ebenfalls für vielfältige Zwecke eingesetzt werden. Dazu gehören vor allem:
 - Die Konstruktion attenuierter Hepatitis C Viren bzw. HCV-ähnlicher Partikel und deren Produktion in Zellkulturen: Durch zufällige oder gezielt hervorgerufene Mutationen, beispielsweise Punktmutationen, Deletionen oder Insertionen, können attenuierte HCV- oder HCV-ähnliche Partikel erzeugt werden, d.h. Viren bzw. virusähnliche Partikel mit voller Replikationskompetenz aber verringerter bzw. fehlender Pathogenität. Solche attenuierte HCV- oder HCV-ähnliche Partikel sind insbesondere als Impfstoff einsetzbar.
 - Die Konstruktion von HCV-RNA-Konstrukten mit integrierten Fremdgenen, beispielsweise zur Verwendung als
 - 35 Leberzell-spezifische Gentherapien in der Genterapie. Auf Grund des ausgeprägten Leberzelltropismus des HCV und der Möglichkeit, Teile des Genoms durch heterologe Sequenzen zu ersetzen, lassen sich HCV-RNA-Konstrukte herstellen, bei denen beispielsweise die Strukturproteine durch ein therapeutisch wirksames Gen ersetzt werden. Das so erhaltene HCV-RNA-Konstrukt wird in Zellen eingeschleust, vorzugsweise mittels Transfektion, die die fehlenden HCV-Funktionen, beispielsweise die Strukturproteine, konstitutiv oder induzierbar exprimieren. Durch diese dem Fachmann unter dem Begriff der 'Transkomplementation' bekannte Technik lassen sich Viruspartikel erzeugen, in die das HCV-RNA-Konstrukt eingebaut wird. Die so erhaltenen Partikel können für die Infektion vorzugsweise von Leberzellen verwendet werden. In diesen wird das therapeutisch wirksame Fremdgen zur Expression gebracht und entfaltet damit seine therapeutische Wirkung.
 - Das Auffinden permissiver Zellen, d.h. Zellen, in denen eine produktive Virusvermehrung erfolgt. Zu diesem Zweck wird entweder eines der vorgenannten HCV-RNA-Genomkonstrukte verwendet, das zur Bildung kompletter infektiöser Viren befähigt ist, oder es wird eines der vorgenannten HCV-Subgenom-Konstrukte eingesetzt, das allerdings zunächst gemäß vorgenanntem Beispiel in eine Zelllinie transfiziert wird, die die fehlenden Funktionen konstitutiv oder induzierbar exprimiert. In all diesen Fällen entstehen Viruspartikel, die zusätzlich zur HCV-Sequenz ein Resistenz- und/oder Reportergen tragen. Zum Auffinden von Zellen, in denen das HCV replizieren kann, werden diese Zellen mit den so hergestellten Viren infiziert und einer Antibiotikumsselektion unterzogen oder,
 - 50 in Abhängigkeit vom HCV-RNA-Konstrukt, mittels Nachweis der Expression des Reportergens untersucht. Da eine Antibiotikumsresistenz bzw. eine Expression des Reportergens nur dann nachweisbar ist, wenn das HCV-RNA-Konstrukt repliziert, müssen die so gefundenen Zellen permissiv sein. Auf diese Weise lassen sich nahezu beliebige Zelllinien oder primäre Zellkulturen hinsichtlich der Permissivität testen und auffinden.
- 55 **[0034]** Das erfindungsgemäße Zellsystem erlaubt auch das gezielte Auffinden von HCV-RNA-Konstrukten, bei denen es auf Grund von Mutationen, die sich entweder zufällig im Rahmen der HCV-RNA-Replikation ereignen oder die gezielt in das Konstrukt eingeführt werden, zu einer Steigerung der Replikationseffizienz kommt. Solche Mutationen, die zu einer Veränderung der Replikation des HCV-RNA-Konstrukts führen, sind dem Fachmann als adaptive

Mutationen bekannt. Die Erfindung umfaßt deshalb auch Verfahren zur Gewinnung von zellkultur-adaptierten Mutanten eines erfindungsgemäßen HCV-RNA-Konstrukts gemäß vorstehender Beschreibung, wobei die Mutanten gegenüber dem originären HCV-RNA-Konstrukt eine erhöhte Replikationseffizienz aufweisen. Sie umfaßt desweiteren ein Verfahren zur Herstellung von Mutanten eines HCV-RNA-Vollängengenoms oder eines HCV-RNA-Teilgenoms oder eines beliebigen HCV-RNA-Konstrukts mit im Vergleich zu dem ursprünglichen HCV-RNA-Vollängengenom oder -Teilgenom oder HCV-RNA-Konstrukt erhöhter Replikationseffizienz, sowie zellkultur-adaptierte Mutanten von HCV-RNA-Konstrukten, HCV-Vollängengenomen und HCV-Teilgenomen mit im Vergleich zu den ursprünglichen Konstrukten, Teil- oder Vollängengenomen erhöhter Replikationseffizienz.

[0035] Das erfindungsgemäße Verfahren zur Gewinnung von zellkultur-adaptierten Mutanten eines erfindungsgemäßen HCV-RNA-Konstrukts, wobei die Mutanten gegenüber dem HCV-RNA-Konstrukt eine erhöhte Replikationseffizienz aufweisen, ist dadurch gekennzeichnet, daß man ein Zellsystem gemäß Anspruch 1, bei dem das eingeschleuste HCV-spezifische Genmaterial ein HCV-RNA-Konstrukt mit Selektionsgen nach einem der Ansprüche 4 bis 19 ist, auf/in dem dem Selektionsgen entsprechenden Selektionsmedium kultiviert, daß man die gewachsenen Zellklone erntet, und daß man aus diesen Zellklonen die HCV-RNA-Konstrukte isoliert.

[0036] Bei einer vorteilhaften Weiterbildung dieses Herstellungsverfahrens werden die isolierten HCV-RNA-Konstrukte wenigstens einmal erneut passagiert, nämlich in Zellen eines Zellsystems nach Anspruch 1 eingeschleust, das dabei erhaltene Zellsystem gemäß Anspruch 1, bei dem das eingeschleuste HCV-spezifische Genmaterial das isolierte HCV-RNA-Konstrukt mit Selektionsgen ist, auf/in dem dem Selektionsgen entsprechenden Selektionsmedium kultiviert, die gewachsenen Zellklone geerntet und daraus die HCV-RNA-Konstrukte isoliert.

Mit dieser Verfahrensvariante kann der Grad der adaptiven Mutationen und damit der Grad der Replikationseffizienz in den betreffenden HCV-RNA-Konstrukten noch gesteigert werden.

[0037] Das erfindungsgemäße Verfahren zur Herstellung von Mutanten eines HCV-Vollängengenoms oder eines HCV-Teilgenoms oder eines beliebigen HCV-RNA-Konstrukts mit im Vergleich zu dem ursprünglichen HCV-Vollängengenom oder -Teilgenom oder HCV-RNA-Konstrukt erhöhter Replikationseffizienz zeichnet sich dadurch aus, daß man mit Hilfe eines der beiden vorstehend genannten Herstellungsverfahren eine zellkultur-adaptierte Variante eines HCV-RNA-Konstrukts herstellt, diese aus den Zellen isoliert, mit im Stand der Technik bekannten Methoden kloniert und sequenziert und durch Vergleich mit der Nukleotid- und Aminosäuresequenz des ursprünglichen HCV-RNA-Konstrukts die Art, Anzahl und Positionen der Mutationen bestimmt, und diese Mutationen dann entweder durch gezielte Mutagenese oder durch Austausch von Sequenzabschnitten, welche die betreffenden Mutationen enthalten, in ein (isoliertes) HCV-Vollängen- oder -teilgenom oder ein beliebiges HCV-RNA-Konstrukt einführt.

Zum Nachweis bzw. zur Verifizierung derjenigen Mutationen, die tatsächlich eine Veränderung der Replikation und insbesondere eine Replikationssteigerung bewirken, kann ein Test durchgeführt werden, bei dem die bestimmten Nukleotid- und/oder Aminosäureaustausche in das ursprüngliche HCV-RNA-Konstrukt eingeführt und dieses wiederum in Zellkultur eingeschleust wird. Wenn die eingeführte Mutation tatsächlich zu einer Steigerung der Replikation führt, sollte im Fall eines HCV-RNA-Konstrukts mit selektierbarem Markergen die Zahl der resistenten Zellklone bei dem künstlich mutierten Konstrukt deutlich höher sein als bei dem unbehandelten Konstrukt. Im Fall eines Konstrukts mit einem Reportergen sollte die Aktivität bzw. Menge des Reporters bei dem künstlich mutierten Konstrukt deutlich höher sein als bei dem unbehandelten.

[0038] Die erfindungsgemäßen zellkultur-adaptierten HCV-RNA-Konstrukte mit hoher Replikationseffizienz sind dadurch gekennzeichnet, daß sie durch Nukleotid- und/oder Aminosäureaustausche von einem HCV-RNA-Konstrukt nach einem der Ansprüche 4 bis 19 ableitbar sind und daß sie mit einem der beiden vorstehend genannten Herstellungsverfahren erhältlich sind.

[0039] Diese zellkultur-adaptierten HCV-RNA-Konstrukte können dazu verwendet werden, beliebige HCV-RNA-Konstrukte oder HCV-Vollängen- oder Teilgenome mit erhöhter Replikationseffizienz herzustellen. Dabei können sowohl Konstrukte mit einem selektierbaren Resistenzgen als auch Konstrukte ohne ein solches bzw. mit einem nicht-selektierbaren Reportergen (z.B. Luziferase) hergestellt werden, denn aufgrund der sehr hohen Replikationseffizienz des zellkultur-adaptierten HCV-RNA-Konstrukts kann dessen Replikation auch in nicht-selektierten Zellen nachgewiesen werden.

Die erfindungsgemäßen zellkultur-adaptierten Mutanten eines HCV-RNA-Konstrukts oder eines HCV-Vollängengenoms oder eines HCV-Teilgenoms mit im Vergleich zu dem ursprünglichen HCV-RNA-Konstrukt oder dem ursprünglichen HCV-Vollängengenom erhöhter Replikationseffizienz sind dadurch charakterisiert, daß sie mit einem Verfahren erhältlich sind, bei dem man in einem zellkultur-adaptierten HCV-RNA-Konstrukt durch Sequenzanalyse und Sequenzvergleich die Art und Anzahl der Mutationen bestimmt und diese Mutationen in ein HCV-RNA-Konstrukt, insbesondere in ein HCV-RNA-Konstrukt gemäß einem der Ansprüche 4 bis 19, oder in ein (isoliertes) HCV-RNA-Vollängengenom einführt, entweder durch gezielte Mutagenese oder durch Austausch von Sequenzabschnitten, die die betreffenden Mutationen enthalten.

[0040] Eine Gruppe ganz bevorzugter HCV-RNA-Konstrukte, HCV-Vollängengenome und HCV-Teilgenomen mit hoher und sehr hoher Replikationseffizienz und infolgedessen sehr guter Eignung für die praktische Anwendung ist

EP 1 043 399 A2

dadurch gekennzeichnet, daß sie einen oder mehrere oder alle der in Tabelle 3 aufgelisteten Aminosäure- bzw. Nukleotidaustausche und/oder einen oder mehrere der folgenden Aminosäureaustausche aufweist: 1283 arg -> gly , 1383 glu -> ala , 1577 lys -> arg , 1609 lys -> glu , 1936 pro -> ser , 2163 glu -> gly , 2330 lys -> glu , 2442 ile -> val. (Die Zahlen beziehen sich auf die Aminosäurepositionen des Polyproteins des HCV-Isolats con1, siehe Tabelle 1).

Besondere Eigenschaften der in den Sequenzprotokollen angegebenen Sequenzen:

SEQ ID-NO: 1

[0041]

Name: I389/Core-3'/wt

Aufbau (Nukleotidpositionen):

1. 1-341: HCV 5' nicht-translatierte Region
2. 342-1193: HCV Core Protein-Neomycin Phosphotransferase Fusionsprotein; selektionierbarer Marker
3. 1202-1812: Interne Ribosomenbindungsstelle des Encephalomyokarditis Virus; erlaubt die Translation des dahinterliegenden HCV offenen Leserasters
4. 1813-10842: HCV Polyprotein von Core bis Nichtstrukturprotein 5B
5. 1813-2385: HCV Core Protein; Strukturprotein
6. 2386-2961: Hüllprotein 1 (envelope protein 1); Strukturprotein
7. 2962-4050: Hüllprotein 2 (envelope protein 2); Strukturprotein
8. 4051-4239: Protein p7
9. 4240-4890: Nichtstrukturprotein 2 (NS2); HCV NS2-3 Protease
10. 4891-6783: Nichtstrukturprotein 3 (NS3); HCV NS3 Protease/Helikase
11. 6784-6945: Nichtstrukturprotein 4A (NS4A); NS3 Protease Kofaktor
12. 6946-7728: Nichtstrukturprotein 4B (NS4B)
13. 7729-9069: Nichtstrukturprotein 5A (NS5A)
14. 9070-10842: Nichtstrukturprotein 5B (NS5B); RNA-abhängige RNA-Polymerase
15. 10846-11076: HCV 3' nicht-translatierte Region

SEQ ID-NO: 2

[0042]

Name: I337/NS2-3'/wt

Aufbau (Nukleotidpositionen):

1. 1-341: HCV 5' nicht-translatierte Region
2. 342-1181: HCV Core Protein-Neomycin Phosphotransferase Fusionsprotein; selektionierbarer Marker
3. 1190-1800: Interne Ribosomenbindungsstelle des Encephalomyokarditis Virus; erlaubt die Translation des dahinterliegenden HCV offenen Leserasters
4. 1801-8403: HCV Polyprotein von Nichtstrukturprotein 2 bis Nichtstrukturprotein 5B
5. 1801-2451: Nichtstrukturprotein 2 (NS2); HCV NS2-3 Protease
6. 2452-4344: Nichtstrukturprotein 3 (NS3); HCV NS3 Protease/Helikase
7. 4345-4506: Nichtstrukturprotein 4A (NS4A); NS3 Protease Kofaktor
8. 4507-5289: Nichtstrukturprotein 4B (NS4B)
9. 5290-6630: Nichtstrukturprotein 5A (NS5A)
10. 6631-8403: Nichtstrukturprotein 5B (NS5B); RNA-abhängige RNA-Polymerase
11. 8407-8637: HCV 3' nicht-translatierte Region

SEQ ID-NO: 3

[0043]

5 Name: I389/NS3-3'/wt

Aufbau (Nukleotidpositionen):

1. 1-341: HCV 5' nicht-translatierte Region
2. 342-1193: HCV Core Protein-Neomycin Phosphotransferase Fusionsprotein; selektionierbarer Marker
- 10 3. 1202-1812: Interne Ribosomenbindungsstelle des Encephalomyokarditis Virus; erlaubt die Translation des dahinterliegenden HCV offenen Leserasters
4. 1813-7767: HCV Polyprotein von Nichtstrukturprotein 3 bis Nichtstrukturprotein 5B
5. 1813-3708: Nichtstrukturprotein 3 (NS3); HCV NS3 Protease/Helikase
- 15 6. 3709-3870: Nichtstrukturprotein 4A (NS4A); NS3 Protease Kofaktor
7. 3871-4653: Nichtstrukturprotein 4B (NS4B)
8. 4654-5994: Nichtstrukturprotein 5A (NS5A)
9. 5995-7767: Nichtstrukturprotein 5B (NS5B); RNA-abhängige RNA-Polymerase
10. 7771-8001: HCV 3' nicht-translatierte Region

20

SEQ ID-NO: 4

[0044]

25 Name: I337/NS3-3'/wt

Aufbau (Nukleotidpositionen):

1. 1-341: HCV 5' nicht-translatierte Region
2. 342-1181: HCV Core Protein-Neomycin Phosphotransferase Fusionsprotein; selektionierbarer Marker
- 30 3. 1190-1800: Interne Ribosomenbindungsstelle des Encephalomyokarditis Virus; erlaubt die Translation des dahinterliegenden HCV offenen Leserasters
4. 1801-7758: HCV Polyprotein von Nichtstrukturprotein 3 bis Nichtstrukturprotein 5B
5. 1801-3696: Nichtstrukturprotein 3 (NS3); HCV NS3 Protease/Helikase
- 35 6. 3697-3858: Nichtstrukturprotein 4A (NS4A); NS3 Protease Kofaktor
7. 3859-4841: Nichtstrukturprotein 4B (NS4B)
8. 4842-5982: Nichtstrukturprotein 5A (NS5A)
9. 5983-7755: Nichtstrukturprotein 5B (NS5B); RNA-abhängige RNA-Polymerase
10. 7759-7989: HCV 3' nicht-translatierte Region

40

SEQ ID-NO: 5

[0045]

45 Name: I389/NS2-3'/wt

Aufbau (Nukleotidpositionen):

1. 1-341: HCV 5' nicht-translatierte Region
2. 342-1193: HCV Core Protein-Neomycin Phosphotransferase Fusionsprotein; selektionierbarer Marker
- 50 3. 1202-1812: Interne Ribosomenbindungsstelle des Encephalomyokarditis Virus; erlaubt die Translation des dahinterliegenden HCV offenen Leserasters
4. 1813-8418: HCV Polyprotein von Nichtstrukturprotein 2 bis Nichtstrukturprotein 5B
5. 1813-2463: Nichtstrukturprotein 2 (NS2); HCV NS2-3 Protease
- 55 6. 2464-4356: Nichtstrukturprotein 3 (NS3); HCV NS3 Protease/Helikase
7. 4357-4518: Nichtstrukturprotein 4A (NS4A); NS3 Protease Kofaktor
8. 4519-5301: Nichtstrukturprotein 4B (NS4B)
9. 5302-6642: Nichtstrukturprotein 5A (NS5A)

EP 1 043 399 A2

10. 6643-8415: Nichtstrukturprotein 5B (NS5B); RNA-abhängige RNA-Polymerase
11. 8419-8649: HCV 3' nicht-translatierte Region

SEQ ID-NO: 6

[0046]

Name: J389/NS3-3/9-13F

Aufbau (Nukleotidpositionen):

1. 1-341: HCV 5' nicht-translatierte Region
2. 342-1193: HCV Core Protein-Neomycin Phosphotransferase Fusionsprotein; selektionierbarer Marker
3. 1202-1812: Interne Ribosomenbindungsstelle des Encephalomyokarditis Virus; erlaubt die Translation des dahinterliegenden HCV offenen Leserasters
4. 1813-7767: HCV Polyprotein von Nichtstrukturprotein 3 bis Nichtstrukturprotein 5B der zellkultur-adaptierten Mutante 9-13F
5. 1813-3708: Nichtstrukturprotein 3 (NS3); HCV NS3 Protease/Helikase
6. 3709-3870: Nichtstrukturprotein 4A (NS4A); NS3 Protease Kofaktor
7. 3871-4653: Nichtstrukturprotein 4B (NS4B)
8. 4654-5994: Nichtstrukturprotein 5A (NS5A)
9. 5995-7767: Nichtstrukturprotein 5B (NS5B); RNA-abhängige RNA-Polymerase 7771-8001: HCV 3' nicht-translatierte Region

SEQ ID-NO: 7

[0047]

Name: J389/Core-3/9-13F

Aufbau (Nukleotidpositionen):

1. 1-341: HCV 5' nicht-translatierte Region
2. 342-1193: HCV Core Protein-Neomycin Phosphotransferase Fusionsprotein; selektionierbarer Marker
3. 1202-1812: Interne Ribosomenbindungsstelle des Encephalomyokarditis Virus; erlaubt die Translation des dahinterliegenden HCV offenen Leserasters
4. 1813-10842: HCV Polyprotein von Core bis Nichtstrukturprotein 5B der zellkultur-adaptierten Mutante 9-13F
5. 1813-2385: HCV Core Protein; Strukturprotein
6. 2386-2961: Hüllprotein 1 (envelope protein 1); Strukturprotein
7. 2962-4050: Hüllprotein 2 (envelope protein 2); Strukturprotein
8. 4051-4239: Protein p7
9. 4240-4890: Nichtstrukturprotein 2 (NS2); HCV NS2-3 Protease
10. 4891-6783: Nichtstrukturprotein 3 (NS3); HCV NS3 Protease/Helikase
11. 6784-6945: Nichtstrukturprotein 4A (NS4A); NS3 Protease Kofaktor
12. 6946-7728: Nichtstrukturprotein 4B (NS4B)
13. 7729-9069: Nichtstrukturprotein 5A (NS5A)
14. 9070-10842: Nichtstrukturprotein 5B (NS5B); RNA-abhängige RNA-Polymerase
15. 10846-11076: HCV 3' nicht-translatierte Region

SEQ ID-NO: 8

[0048]

Name: J389/NS3-3/5.1

Aufbau (Nukleotidpositionen):

1. 1-341: HCV 5' nicht-translatierte Region
2. 342-1193: HCV Core Protein-Neomycin Phosphotransferase Fusionsprotein; selektionierbarer Marker

3. 1202-1812: Interne Ribosomenbindungsstelle des Encephalomyokarditis Virus; erlaubt die Translation des dahinterliegenden HCV offenen Leserasters
 4. 1813-7767: HCV Polyprotein von Nichtstrukturprotein 3 bis Nichtstrukturprotein 5B der zellkultur-adaptierten Mutante 5.1

5

5. 1813-3708: Nichtstrukturprotein 3 (NS3); HCV NS3 Protease/Helikase
 6. 3709-3870: Nichtstrukturprotein 4A (NS4A); NS3 Protease Kofaktor
 7. 3871-4653: Nichtstrukturprotein 4B (NS4B)
 8. 4654-5994: Nichtstrukturprotein 5B (NS5A)

10

9. 5995-7767: Nichtstrukturprotein 5B (NS5B); RNA-abhängige RNA-Polymerase 7771-8001: HCV 3' nicht-translatierte Region

SEQ ID-NO: 9

15 [0049]

Name: |389/Core-3/5.1

Aufbau (Nukleotidpositionen):

20

1. 1-341: HCV 5' nicht-translatierte Region
 2. 342-1193: HCV Core Protein-Neomycin Phosphotransferase Fusionsprotein; selektionierbarer Marker
 3. 1202-1812: Interne Ribosomenbindungsstelle des Encephalomyokarditis Virus; erlaubt die Translation des dahinterliegenden HCV offenen Leserasters
 4. 1813-10842: HCV Polyprotein von Core bis Nichtstrukturprotein 5B der zellkultur-adaptierten Mutante 5.1

25

5. 1813-2385: HCV Core Protein; Strukturprotein
 6. 2386-2961: Hüllprotein 1 (envelope protein 1); Strukturprotein
 7. 2962-4050: Hüllprotein 2 (envelope protein 2); Strukturprotein
 8. 4051-4239: Protein p7
 9. 4240-4890: Nichtstrukturprotein 2 (NS2); HCV NS2-3 Protease
 10. 4891-6783: Nichtstrukturprotein 3 (NS3); HCV NS3 Protease/Helikase
 11. 6784-6945: Nichtstrukturprotein 4A (NS4A); NS3 Protease Kofaktor
 12. 6946-7728: Nichtstrukturprotein 4B (NS4B)
 13. 7729-9069: Nichtstrukturprotein 5A (NS5A)
 14. 9070-10842: Nichtstrukturprotein 5B (NS5B); RNA-abhängige RNA-Polymerase
 15. 10846-11076: HCV 3' nicht-translatierte Region

30

35

SEQ ID-NO: 10

40 [0050]

Name: |389/NS3-3/19

Aufbau (Nukleotidpositionen):

45

1. 1-341: HCV 5' nicht-translatierte Region
 2. 342-1193: HCV Core Protein-Neomycin Phosphotransferase Fusionsprotein; selektionierbarer Marker
 3. 1202-1812: Interne Ribosomenbindungsstelle des Encephalomyokarditis Virus; erlaubt die Translation des dahinterliegenden HCV offenen Leserasters
 4. 1813-7767: HCV Polyprotein von Nichtstrukturprotein 3 bis Nichtstrukturprotein 5B der zellkultur-adaptierten Mutante 19

50

5. 1813-3708: Nichtstrukturprotein 3 (NS3); HCV NS3 Protease/Helikase
 6. 3709-3870: Nichtstrukturprotein 4A (NS4A); NS3 Protease Kofaktor
 7. 3871-4653: Nichtstrukturprotein 4B (NS4B)
 8. 4654-5994: Nichtstrukturprotein 5A (NS5A)

55

9. 5995-7767: Nichtstrukturprotein 5B (NS5B); RNA-abhängige RNA-Polymerase 7771-8001: HCV 3' nicht-translatierte Region

SEQ ID-NO: 11

[0051]

5 Name: I389/Core-3/19
Aufbau (Nukleotidpositionen):

1. 1-341: HCV 5' nicht-translatierte Region
2. 342-1193: HCV Core Protein-Neomycin Phosphotransferase Fusionsprotein; selektionierbarer Marker
- 10 3. 1202-1812: Interne Ribosomenbindungsstelle des Encephalomyokarditis Virus; erlaubt die Translation des dahinterliegenden HCV offenen Leserasters
4. 1813-10842: HCV Polyprotein von Core bis Nichtstrukturprotein 5B der zellkultur-adaptierten Mutante 19
5. 1813-2385: HCV Core Protein; Strukturprotein
- 15 6. 2386-2961: Hüllprotein 1 (envelope protein 1); Strukturprotein
7. 2962-4050: Hüllprotein 2 (envelope protein 2); Strukturprotein
8. 4051-4239: Protein p7
9. 4240-4890: Nichtstrukturprotein 2 (NS2); HCV NS2-3 Protease
10. 4891-6783: Nichtstrukturprotein 3 (NS3); HCV NS3 Protease/Helikase
- 20 11. 6784-6945: Nichtstrukturprotein 4A (NS4A); NS3 Protease Kofaktor
12. 6946-7728: Nichtstrukturprotein 4B (NS4B)
13. 7729-9069: Nichtstrukturprotein 5A (NS5A)
14. 9070-10842: Nichtstrukturprotein 5B (NS5B); RNA-abhängige RNA-Polymerase
15. 10846-11076: HCV 3' nicht-translatierte Region

25 [0052] Die Erfindung wird im folgenden anhand von Ausführungsbeispielen und dazugehörigen Tabellen und Figuren näher erläutert. Die erwähnten Figuren zeigen

Fig. 1 A: Die Struktur eines erfindungsgemäßen HCV-RNA-Konstrukts Ganz oben ist eine schematische Darstellung der Struktur des kompletten parentalen HCV-Genoms gegeben mit den Positionen der Gene für die Spaltungsprodukte core, E1, E2, p7, NS2, NS3, NS4A, NS4B, NS5A und NS5B innerhalb des Polyproteins, und den 5' und 3' nichttranslatierten Regionen (5' NTR und 3' NTR) — als Horizontalbalken dargestellt —, und mit den beiden für die Erzeugung der Subgenom-Konstrukte ausgewählten Positionen, nämlich der Position der 'GDD-katalytischen Domäne' der NS5B RNA Polymerase (GDD) und der Position der 3' Grenze der HCV-IRES (Nukleotidpositionen 1 bis 377 bzw. 1 bis 389) — oberhalb des Genomschemas eingezeichnet —. Die Zahlen unterhalb des Genomschemas bezeichnen die entsprechenden Nukleotidpositionen.

Darunter sind schematische Darstellungen der Strukturen zweier erfindungsgemäßer, modifizierter HCV-RNA-Konstrukte (Subgenom) gezeigt, bestehend aus der 5' HCV-IRES, dem Neomycinphosphotransferasegen (Neo^r), der EMCV-IRES (E-I) und den HCV Sequenzen von NS2 bzw. NS3 bis zum authentischen 3' Ende. Die Position der das NS5B Polymerase GDD-Motiv umfassenden 10 Aminosäuren-Deletion ist jeweils mit einem Dreieck (Δ) markiert.

Fig. 1 B: Das Ergebnis einer denaturierenden Formaldehyd-Agarose-Gelelektrophorese zum Nachweis von replizierter Plusstrang-RNA in transfizierten subpassagierten Huh-7 Zellklonen. Die Positionen der HCV-spezifischen RNAs (Pfeile) und der 28S rRNA sind rechts von Spur 12 angegeben, die Größen (Anzahlen der Nukleotide) der RNA-Marker (M) sind links von Spur 1 angegeben.

Fig. 1 C : Das Ergebnis eines PCR-Tests mit nachfolgendem Southern-Blot zum Nachweis der Abwesenheit von integrierter Replikon-DNA in den meisten der selektierten Zellklone. Spuren 1 und 2 zeigen die Positivkontrollen, Spur 13 die Negativ-Kontrolle. Die Zahlenangaben links der Spur 1 bezeichnen die Größe der Nukleotid-Marker-Moleküle.

Fig. 2 A: Das Ergebnis eines PCR-Tests mit nachfolgendem Southern-Blot zum sensitiven Ausschluss integrierter Replikon-DNA (Plasmid Moleküle I₃₇₇/NS3-3'/wt) in einem HCV-RNA-Konstrukt-haltigen Zellklon (9-13). Die Spuren 7 bis 11 repräsentieren das Ergebnis einer Titration von DNA-Molekülen des Konstrukts I₃₇₇/NS3-3'/wt ohne Zugabe von Gesamt-DNA des Zellklons 9-13 und die Spuren 2 - 6 repräsentieren die gleichen Plasmidmoleküle mit Zugabe von jeweils 1 µg 9-13 DNA vor der PCR (zwecks Ausschluss

eines Inhibitors der PCR in der DNA-Präparation). Spur 13 repräsentiert die Negativ-Kontrolle (PCR ohne DNA-Sonde). Spur 1 zeigt das Ergebnis, das mit einem μg Gesamt-DNA des Zellkorns 9-13 erhalten wurde.

- 5 **Fig. 2 B:** Das Ergebnis eines Northern-Blot-Tests zur Quantifizierung von HCV Plus- und Minusstrang RNA. Die Pfeile markieren die Positionen von Replikon-RNA. Die "plus" und "minus" -Abgaben bezeichnen die positive (plus) bzw. negative (minus) Polarität der RNA-Kontrollen, die auf das Gel aufgetragen wurden. "Minusstrang" und "Plusstrang" bezeichnen die Spezifität der radioaktiven RNA-Sonden.
- 10 **Fig. 2 C:** Ergebnis einer Formaldehyd-Agarose-Gelelektrophorese nach radioaktiver Markierung der intrazellulär replizierten HCV-RNA zum Nachweis der Resistenz der HCV-RNA-Replikation gegen Dactinomycin.
- Fig. 3 A:** Nachweis von HCV-spezifischen Antigenen in den selektierten Zellklonen mittels Immunopräzipitation nach metabolischer Radioaktivmarkierung. Die Spuren 7 - 9 repräsentieren authentische Größenmarker (die nach transienter Expression eines HCV-RNA-Konstrukts in Huh-7-Zellen erhalten wurden); identifizierte HCV-Proteine sind am linken Rand von Spur 1 markiert, die Molekulargewichte (in Kilodalton) sind am rechten Rand von Spur 9 angegeben.
- 15 **Fig. 3 B:** Ergebnisse eines Immunfluoreszenztests zum Nachweis der subzellulären Lokalisation von HCV Antigenen.
- 20 **Fig. 4:** Schematische Darstellung der Struktur eines erfindungsgemäßen selektierbaren HCV-RNA-Konstrukts (komplettes Genom) bestehend aus der 5' HCV-IRES, dem Neomycinphosphotransferasagen (NeoR), einem heterologen IRES-Element, z.B. des Encephalomyocarditisvirus (E-I), dem vollständigen HCV-Leseraster und der authentischen 3' NTR.
- 25 **Fig. 5:** Schematische Darstellung der Struktur von HCV-RNA-Konstrukten mit inseriertem Antibiotikumresistenzgen (A) innerhalb der für das Polyprotein kodierenden Nukleotidsequenz (monocistronische RNA), und (B) innerhalb der 3' NTR (bicistronische RNA).
- 30 **Fig. 6:** Schematische Darstellung der Struktur von HCV-RNA-Konstrukten mit inseriertem Reportergen (A) als Teil eines HCV-Replikons von NS3 bis NS5B; — das Reportergen wird letztendlich durch virale oder durch zelluläre Proteasen aus dem Polyprotein gespalten und das selektierbare Markergen (Selektionsgen) bzw. das Resistenzgen durch Kontransfektion in die Zellen einschleust, (B) als Teil eines Fusionsgens aus Resistenz- und Reportergen (z.B. für die Neomycinphosphotransferase und green fluorescent Protein) (C) als Teil eines Replikons aus Resistenz- und Reportergen (z.B. für die Neomycinphosphotransferase und das green fluorescent Protein), die über eine Nukleotidsequenz verbunden sind, welche für eine Aminosäuresequenz kodiert (schraffierter Bereich), die von einer Protease gespalten werden kann oder die über eine selbstspaltende (autokatalytische) Aktivität verfügt, (D) als unabhängiges Gen (hier green fluorescent protein), das von einer eigenen internen Ribosomenbindungsstelle (IRES) aus exprimiert wird; — das Resistenzgen (hier: Neomycinphosphotransferase-Gen) wird davon unabhängig ebenfalls von einer eigenen internen Ribosomenbindungsstelle (IRES) aus exprimiert (polycistronisches Konstrukt).
- 40 **Fig. 7:** Schematische Darstellung der Struktur eines HCV-RNA-Konstrukts bei dem das Resistenzgen über ein Ribozym bzw. eine Erkennungsstelle für ein Ribozym mit der HCV-RNA-Sequenz verbunden ist. Die dicken Linien stellen die HCV 5' und 3' NTRs dar, E-I ist eine heterologe interne Ribosomenbindungsstelle, die für die Expression des Resistenzgens notwendig ist, und das graue Quadrat stellt das Ribozym bzw. eine Erkennungsstelle für ein Ribozym dar.
- 45 **Fig. 8:** Schematische Darstellung der Struktur eines HCV-RNA-Konstrukts mit Resistenzgen und integriertem Fremdgen.
- 50 **Fig. 9:** Methodisches Vorgehen zum Vergleich der spezifischen Infektiosität (ausgedrückt als Anzahl gebildeter Zellkolonien) von Gesamt-RNA versus in vitro Transkripte. HCV-RNA wird mittels in vitro Transkription eines entsprechenden RNA-Konstrukts hergestellt und durch Messung der optischen Dichte bei 260 nm (OD 260 nm) quantifiziert. Eine definierte Anzahl dieser Moleküle wird mit einer bestimmten Menge Gesamt-RNA von naiven Huh-7 Zellen gemischt und diese Mischung mit Hilfe der Elektroporation in
- 55

naïve Huh-7 Zellen eingeschleust. Parallel dazu wird die Gesamt-RNA eines Zellklons, der mit der in Figur 1 beschriebenen Methode hergestellt wurde, mit einem im Stand der Technik bekannten Verfahren isoliert und die Menge der darin enthaltenen HCV-RNA mittels Northern-blot unter Verwendung einer HCV-spezifischen RNA-Sonde und anschließender Quantifizierung mittels Phosphorimager bestimmt. Eine definierte Menge dieser Gesamt-RNA wird analog den *in vitro* Transkripten in naïve Huh-7 Zellen transfiziert. Diese Zellen in beiden Ansätzen werden danach einer G418-Selektion unterzogen und die Anzahl der gebildeten Kolonien durch Auszählen nach Fixieren und anfärben mit Coomassie-Brilliant-Blau bestimmt. Zur Bestimmung der Transfektionseffizienz wird jedem Transfektionsansatz 1 µg eines Plasmids zugesetzt, das die Expression der Luciferase erlaubt. Ein Aliquot der transfizierten Zellen wird nach 24 Stunden geerntet und die Luciferaseaktivität im jeweiligen Zellysats bestimmt. Die Anzahl der Kolonien wird jeweils auf die Luciferaseexpression normiert.

Fig. 10: Sequenzanalyse der 9-13 Klone. Gesamt-RNA des Zellklons 9-13, der durch Transfektion des HCV-RNA-Konstrukts 1377/NS3-3' entstand, wurde mit einem im Stand der Technik bekannten Verfahren isoliert und das HCV-RNA-Konstrukt von Nukleotidposition 59 bis 9386 mit Hilfe der 'long-distance RT-PCR' unter Verwendung der primer S59 und A9413 amplifiziert. Die PCR-Fragmente wurden kloniert und 11 Klone (genannt 9-13 A - K) vollständig sequenziert, wobei sich die Klone D und I, E und G sowie H und J als identisch erwiesen. Die Positionen der Aminosäureunterschiede in der NS3-5B Region zwischen den rekombinierten HCV-RNAs und dem parentalen Konstrukt sind mit einem dicken vertikalen Strich beim jeweiligen Klon markiert. Jeder Klon wurde mit dem Restriktionsenzym *Sfi* I verdaut und das jeweilige Fragment in das parentale Konstrukt inseriert. Diese Klone wurden jeweils in Huh-7 Zellen transfiziert und die Zellen wie in Figur 1 beschrieben einer Selektion unterzogen. Die Anzahl der mit jedem Konstrukt erhaltenen Zellklone ist rechts neben dem jeweiligen Konstrukt vermerkt.

Fig. 11 A: Prinzip der Replikationsbestimmung mit Hilfe eines Reportergens. Im oberen Teil der Figur ist das HCV-DNA-Konstrukt 1389/Luc/NS3-3' dargestellt, bestehend aus der HCV 5' NTR (Nukleotidposition 1-389), dem Luciferasegen (*luc*), der IRES des Encephalomyocarditis Virus, dem HCV NS3-5B und der 3' NTR. Die Position des aktiven Zentrums der NSB RNA-Polymerase, in das ein inaktivierender Aminosäureaustausch eingeführt wurde, ist mit 'GND' angegeben. Die Plasmide, die für das replikationskompetente bzw. das defekte HCV-RNA-Konstrukt kodieren, werden mit dem Restriktionsenzym *Sca* I verdaut und in eine *in vitro* Transkription mit der T7 RNA-Polymerase eingesetzt. Nach Entfernung der Matrizen-DNA werden die jeweiligen HCV-RNA-Konstrukte mittels Elektroporation in naïve Huh-7 Zellen eingeschleust und diese in regelmäßigen Abständen geerntet.

Fig. 11 B: Vergleich der Luciferaseaktivitäten in Zellen transfiziert mit dem parentalen HCV-RNA-Konstrukt 1389/Luc/NS3-3'/wt (wt) oder den folgenden Varianten: Der inaktiven RNA (318 DN), der Variante 9-13F oder der Variante 5.1. Die Zellen wurden 6 (nicht gezeigt), 24, 48, 72, 96, 120, 144 und 168 Stunden nach der Transfektion geerntet und die Luciferaseaktivitäten luminometrisch bestimmt.

Fig. 12: Selektionierbare HCV-Vollängengenome (Konstrukte 1389/core-3'/5.1 und 1389/core-3'/9-13F).

(A) Schematische Darstellung des Vollängerkonstrukts. Der Bereich zwischen den beiden angegebenen Erkennungsstellen für das Restriktionsenzym *Sfi* I entspricht den Sequenzen der hochadaptierten RNA-Varianten 5.1. oder 9-13F.

(B) Anzahl der Kolonien die nach Transfektion von jeweils 0,1 µg *in vitro* transkribierter RNA der unter A dargestellten Konstrukte 1389/core-3'/5.1 in HUH7-Zellen erhalten wurden. Angegeben ist das Ergebnis eines repräsentativen Experimentes.

(C) Nachweis autonom replizierender HCV-Vollängen-RNAs in G418-resistenten Zellklonen, die nach Transfektion des entsprechenden *in vitro* Transkripts erhalten wurden. Die Abbildung zeigt das Autoradiogramm eines Northern Blots, der mit einer Sonde gegen das *neo*-Resistenzgen und der HCV 5' NTR hybridisiert wurde. Die in Spur 1 und 2 dargestellten Kontrollen entsprechen jeweils 10⁶ Molekülen der angegebenen *in vitro* Transkripte, gemischt mit Gesamt-RNA aus naiven Huh-7 Zellen. Die Negativkontrolle enthält ausschließlich Gesamt RNA aus naiven Huh-7 Zellen (Spur 3). Die Spuren 4-9 enthalten 3-10 µg Gesamt-RNA aus G418-resistenten Zellklonen, die nach Transfektion von *in vitro* transkribierter 1389/core-3'/5.1-RNA bzw. 1389/core-3'/9-13F-RNA erhalten wurden. Die für die Selektion verwendete G418-Konzentration ist jeweils angegeben. Fünf der dargestellten Zellklone enthalten die hoch adaptierte RNA-Variante 5.1 (Spur 4-8), einer die adaptierte RNA-Variante 9-13F (Spur 9).

Fig. 13: HCV-RNA-Konstrukte mit einem Reportergen. (A) Bicistronische HCV-RNA-Konstrukte. Das Reportergen wird mit Hilfe einer separaten IRES translatiert. (B) Monocistronische HCV-RNA-Konstrukte. Das Reportergenprodukt wird als Fusionsprotein mit einem HCV-Protein exprimiert. Die beiden Anteile sind über eine Erkennungssequenz für eine virale oder zelluläre Protease verbunden, die eine proteolytische Trennung der beiden fusionierten Proteinanteile erlaubt. Im gezeigten Beispiel wurden das Reportergenprodukt und das jeweilige HCV-Protein über eine Erkennungssequenz für Ubiquitin (Ub) fusioniert.

Fig. 14: Tricistronisches Vollängen HCV-RNA-Konstrukt, das zusätzlich zum Resistenzgen ein Fremdgen inseriert besitzt.

Fig. 15: Monocistronische HCV-RNA-Konstrukte, bei denen das Resistenzgenprodukt als Fusionsprotein mit dem HCV-Anteil exprimiert wird. Das Resistenzgen (RG) ist entweder als Fusionsprotein aktiv oder es wird so mit einer proteolytisch spaltbaren Sequenz mit dem HCV-Anteil fusioniert, daß das Resistenzgenprodukt durch eine zelluläre oder virale Protease vom HCV-Anteil abgespalten wird. Im gezeigten Beispiel wurde das Resistenzgen über die für Ubiquitin (Ub) kodierende Sequenz mit dem jeweiligen HCV-Anteil fusioniert.

Beispiel 1: Herstellung von HCV-RNA-Konstrukten

(A) Synthese und Klonierung eines vollständigen HCV-Konsensusgenoms mittels RT-PCR

[0053] Aus der Leber eines chronisch infizierten Patienten wurde das HCV-Genom, d.h. die HCV-RNA wie nachfolgend beschrieben isoliert.

[0054] Aus ca. 100 mg Leber wurde die komplette RNA gemäß dem Verfahren von Chomczynski und Sacchi (1987, Anal. Biochem. 162, 156) isoliert. Mit 1 µg dieser isolierten RNA wurde eine reverse Transkription mit den Primern A6103 (GCTATCAGCCGGTTTCCTCACTGC) oder A9413 (CAGGATGGCCTATTGG CCTGGAG) und dem "expand reverse transcriptase"-System (Boehringer Mannheim, Deutschland) nach den Vorschriften des Herstellers durchgeführt. Mit den Produkten dieser reversen Transkription (RT) wurde eine Polymerase-Kettenreaktion (PCR/polymerease chain reaction) durchgeführt, und zwar unter Verwendung des "expand low template"-Systems (Boehringer Mannheim, Deutschland), wobei der Puffer mit 2% Dimethylsulfoxid-Gehalt eingesetzt wurde. Nach einer Stunde bei 42°C wurde 1/8 dieses Reaktionsansatzes in einem ersten PCR-Durchgang mit den Primern A6103 und S59 (TGTTCTTCACGCA-GAAAGCGTCTAG) oder A9413 und S4542 (GATGAGCT CGCCGCGAAGCTGTCC) eingesetzt. Nach 40 Zyklen wurde 1/10 dieses Reaktionsansatzes in einem zweiten PCR-Durchgang mit den Primern S59 und A4919 (AGCACAGCCGCGCTCATGCACTCG) oder S4542 und A9386 (TTAGCTCCCCG TTCATCGGTTGG) eingesetzt. Nach 30 Zyklen wurden die PCR-Produkte mittels präparativer Agarose-Gel-Elektrophorese gereinigt und die dabei eluierten Fragmente wurden in den Vektor pCR2.1 (Invitrogen) oder pBSK II (Stratagene) ligiert. Vier Klone von jedem Fragment wurden analysiert und sequenziert, und es wurde eine Konsensus-Sequenz ermittelt. Zu diesem Zweck wurden die DNA-Sequenzen miteinander verglichen. Die Positionen, an denen sich die Sequenz eines der Fragmente von den übrigen unterschied, wurde als unerwünschte Mutation betrachtet. Im Fall von Mehrdeutigkeiten der Sequenz wurden kürzere sich überlappende PCR-Fragmente der betreffenden Region amplifiziert und mehrere Klone sequenziert. Auf diese Weise konnten zahlreiche potentielle Mutationen in jedem Fragment identifiziert und somit eine isolat-spezifische Konsensussequenz etabliert werden. Diese etablierte Konsensussequenz bzw. dieses Genom gehört zum weltweit verbreiteten Genotyp 1b. Die nicht translatierte Region am 3'-Ende (=3' NTR) wurde mittels konventioneller PCR erhalten, wobei ein Antisense-Primer eingesetzt wurde, der die letzten 24 Nukleotide des im Stand der Technik bekannten 'X'-tails (Tanaka et al., 1995, Biochem. Biophys. Res. Commun. 215, 744; und Rice, PCT/US 96/14033) abdeckt. Die authentische nicht translatierte Region am 5'-Ende (=5' NTR) strangabwärts vom T7 Promotor wurde mittels PCR erzeugt, wobei zum einen ein Oligonukleotid verwendet wurde, das einem verkürzten T7 Promotor (TAA TAC GAC TCA CTA TAG) und den ersten 88 Nukleotiden von HCV entspricht, und zum anderen eines der vorgenannten Plasmide eingesetzt wurde, das eines der 5' Fragmente des Genoms trägt. Aus den subgenomischen Fragmenten mit der geringsten Anzahl an Nicht-Konsensus-Austauschen wurde ein komplettes HCV-Konsensusgenom zusammengesetzt und in einen modifizierten pBR322-Vektor inseriert. Abweichungen von der Konsensussequenz wurden mittels ortsgerechter Mutagenese ("site-directed mutagenesis") beseitigt. Um "run-off"-Transkripte mit einem authentischen 3' Ende herzustellen, wurde die 3'-NTR der Isolate (mit dem Ende TGT) zu AGT modifiziert (gemäß der Sequenz vom Genotyp 3 = Klon 'WS' nach Kolykhalov et al., 1996, J. Virol. 70, 3363) und außerdem wurde ein zusätzlicher Nukleotidaustausch an Position 9562 vorgenommen, um die A:T Basenpaarung in der Haarnadelstruktur am 3' Ende der 3' NTR (Kolykhalov et al. bild.) beizubehalten. Um eine interne Restriktionsstelle für das Enzym ScaI zu beseitigen, wurde ferner ein sog. stiller ("silent") Nukleotidaustausch vorgenommen. Nach dem Zusammenfügen des Vollängen-Genoms mit passenden 5'- und 3' NTRen wurde die komplette HCV-Sequenz überprüft. Dabei wurde kein ungewünschter Nukleotidaustausch

gefunden.

[0055] Das auf diese Weise hergestellte HCV-Genom sollte per Definition hepatotrop sein.

(B) Synthese selektierbarer HCV-Subgenom-Konstrukte

5

[0056] Unter Verwendung des unter (A) beschriebenen Konsensusgenoms wurden HCV-Subgenom-Konstrukte hergestellt, die das Antibiotikumsistenzgen Neomycin-Phosphotransferase (NPT) und zwei Sequenzen von internen Ribosomenbindungsstellen (IRES) enthalten. Die hierfür angewendeten biochemischen Verfahrenstechniken sind dem Fachmann bekannt und geläufig (siehe: Sambrook, J., E.F. Fritsch, T. Maniatis, 1989, *Molecularcloning: a laboratory manual*, 2nd ed., Cold Spring Harbour Laboratory, Cold Spring Harbor, N.Y.; Ausubel et al. (eds.), 1994, *Current Protocols in Molecular Biology*, Vol. 1-3, John Wiley & Sons Inc., New York). Das Antibiotikumsistenzgen wurde unmittelbar hinter der 5' NTR inseriert, wodurch eine bicistronische RNA erhalten wurde (siehe Fig. 1 A). Ebenso gut kann das Antibiotikumsistenzgen aber auch an anderer Stelle des HCV-Subgenom-Konstrukts inseriert werden, beispielsweise innerhalb der für das Polyprotein kodierenden Nukleotidsequenz, wodurch eine monocistronische RNA erhalten wird (siehe Fig. 5 A) oder in die 3' NTR (siehe Fig. 5 B). Bei den IRES-Elementen handelt es sich zum einen um eine der beiden HCV-IRES-Varianten Nukleotide 1-377 oder Nukleotide 1-389, und zum anderen um die IRES des Enzephalomyocarditis Virus, die die Translation der HCV Sequenz strangabwärts von den Genen für NS2 oder NS3 bis zu dem authentischen 3' Ende des Genoms steuert.

[0057] Die beiden genannten HCV-IRES-Varianten wurden wie folgt ermittelt: Auf der Basis von Deletionsanalysen der 3' Grenze der HCV-IRES (Reynolds et al. 1995, *EMBO J.* 14, 6010) wurden verschiedene Abschnitte der 5' NTR mit dem NPT Gen fusioniert und anhand von Kotransfektionen mit einem das T7 RNA Polymerase Gen enthaltenden Plasmid hinsichtlich der maximalen Anzahl gebildeter Kolonien analysiert. Die besten Ergebnisse wurden mit den HCV Sequenzen von 1-377 und 1-389 erhalten. Da sich das AUG-Startkodon des HCV Polyproteins an Position 342 befindet und somit in der IRES-Sequenz enthalten ist, kommt es zu einer Fusion von 12 bzw. 16 Aminosäuren des HCV-Kapsidproteins ("Core-Proteins") mit der Neomycin Phosphotransferase (siehe Fig. 1 A).

[0058] Diese modifizierten HCV-Subgenom-Konstrukte erhielten dementsprechend die Bezeichnungen $I_{377}/NS2-3'$ (oder $I_{377}/NS3-3'$) und $I_{389}/NS2-3'$ (oder $I_{389}/NS3-3'$). Sie sind in Fig. 1A schematisch dargestellt.

[0059] Mit in-vitro-Transkripten dieser modifizierten parentalen HCV-Subgenom-Konstrukte $I_{377}/NS2-3'$ (oder $I_{377}/NS3-3'$) und $I_{389}/NS2-3'$ (oder $I_{389}/NS3-3'$) wurden verschiedene Zelllinien und Primärzellkulturen von menschlichen Hepatocyten transfiziert.

[0060] Als parallele Negativ-Kontrolle zu allen Transfektionsexperimenten wurde zu jedem modifizierten parentalen HCV-Subgenom-Konstrukt ein entsprechend modifiziertes aber defektes Subgenom konstruiert, das sich von dem parentalen dadurch unterscheidet, daß es innerhalb des Leserasters eine Deletion von 10 Aminosäuren aufweist, die das aktive Zentrum der NSSB RNA Polymerase umfaßt (Behrens et al., 1996, *EMBO J.* 15, 12; und Lohmann et al., 1997, *J. Virol.* 71, 8416).

(C) Synthese selektierbarer HCV-Genom-Konstrukte

[0061] Ein NS2-3' Subgenomkonstrukt, das am 5' Ende mit einem Fragment des Luziferasegens und der vollständigen EMCV-IRES verbunden ist, wurde mit NcoI und SpeI restringiert und mittels präparativer Agarosegelelektrophorese gereinigt. Der so erhaltene Vektor wurde in einer 3-Faktor Ligation mit einem NcoI/NotI-HCV-Fragment, entsprechend den Nukleotidpositionen 342 bis 1968 des HCV-Genoms und mit einem NotI/SpeI-Fragment, entsprechend den Nukleotidpositionen 1968-9605 ligiert. Das entstandene Konstrukt, bei dem das vollständige HCV-Leseraster und die 3' NTR stromabwärts dem Luziferasegenfragment und der EMCV-IRES liegen, wurde danach mit PmeI und SpeI restringiert und mit dem analog restringierten $I_{389}/NS3-3'/wt$ -Subgenomkonstrukt-Vektor ligiert. Dieses selektionierbare HCV-Genomkonstrukt ist in Fig. 4 dargestellt.

(D) Herstellung von den HCV-RNA-Konstrukten entsprechenden in-vitro-Transkripten

[0062] Die vorstehend beschriebenen gereinigten Plasmid DNAs wurden mit Scal linearisiert und nach Phenol/Chloroform-Extraktion und Isopropanol-Präzipitation in eine in-vitro-Transkriptionsreaktion eingesetzt unter Verwendung der folgenden Komponenten: 80 mM HEPES, pH 7.5, 12,5 mM $MgCl_2$, 2 mM Spermidin, 40 mM Dithiothreitol, 2 mM von jedem NTP, 1 Einheit RNasin/ μ l, 50 μ g/ml restringierte DNA und ca. 2 Einheiten/ μ l T7 RNA Polymerase. Nach 2 Std. bei 37°C wurde die Hälfte der Menge an T7 Polymerase zugegeben und der Reaktionsansatz weitere 2h inkubiert. Zur Entfernung von DNA wurde die Mischung mit saurem Phenol extrahiert (U. Kedzierski, J.C. Porte, 1991, *Bio Techniques* 10, 210), mit Isopropanol präzipitiert, das Pellet in Wasser gelöst und mit DNase (2 Einheiten pro μ g DNA) für 60 Min. bei 37°C inkubiert. Nach anschließender Extraktion mit saurem Phenol, saurem Phenol/Chloroform und Chloroform und Isopropanol-Präzipitation wurde die gelöste RNA mittels optischer Dichtemessungen quantifiziert und

ihre Unversehrtheit mittels Formaldehyd-Agarose-Gelelektrophorese überprüft.

Beispiel 2: Transfektionsexperimente mit der Hepatomazelllinie Huh-7

[0063] Bei sämtlichen Transfektionsexperimenten wurde sorgfältig darauf geachtet, daß jegliche Matrizen-DNA zuvor entfernt worden war, um zu vermeiden, daß solche DNA in transfizierte Zellen integrieren und diesen unabhängig von einer HCV-Replikation eine Neomycin-Resistenz vermitteln konnte. Deshalb wurde im Anschluß an die in-vitro-Transkription (Beispiel 1 D) die Reaktionsmischung mit 2 Einheiten DNase pro µg DNA für 60 Min. bei 37°C behandelt und mit saurem Phenol, saurem Phenol/Chloroform und Chloroform extrahiert. Vor der Verwendung für die Transfektion wurde die präzipitierte RNA mittels Formaldehyd Agarose Gel Elektrophorese analysiert.

[0064] Es wurden drei separate Transfektionsexperimente mit der hoch differenzierten humanen Hepatomazelllinie Huh-7 (gemäß Nakabayashi et al. 1982, *Cancer Res.* 42, 3858) durchgeführt. Dabei wurde jeweils 15 µg RNA in 8 x 10⁵ Huh-7-Zellen mit Hilfe der Elektroporation eingebracht und diese Zellen anschließend in Kulturschalen von 10 cm Durchmesser ausgesät. 24 Stunden nach der Aussaat wurde Neomycin (= G418) in einer Endkonzentration von 1 mg/ml zugegeben. Das Kulturmedium wurde zweimal pro Woche gewechselt. Nach 3 - 5 Wochen waren kleine Kolonien erkennbar, die isoliert und unter den gleichen Kulturbedingungen passagiert wurden.

[0065] Die Zellkline, die im Verlauf des ersten Experiments erhalten wurden, wurden isoliert und subpassagiert. Während dieser Prozedur starben die meisten Klone und die Endausbeute betrug nur noch 9 Klone von Zellen, die mit den parentalen HCV-Subgenom-Konstrukten transfiziert worden waren und 1 Klon (Klone 8-1) von Zellen, die mit einem defekten HCV-Genom-Konstrukt, nämlich einer defekten NS2-3' HCV-RNA transfiziert worden waren. Außer einer verkürzten Verdopplungszeit und dem gelegentlichen Auftreten von unregelmäßig geformten Zellen wurden keine beständigen morphologischen Unterschiede zwischen diesen 9 Zellklonen und dem einen Zellklon (Klone 8-1) oder den parentalen Huh-7 Zellen gefunden.

[0066] Die Hauptkriterien für funktionierende HCV-Genomkonstrukte sind die Bildung von viraler RNA mit korrekter Größe und die Abwesenheit von (integrierter) Plasmid DNA, die eine G418-Resistenz übertragen bzw. vermitteln könnte.

[0067] Um die HCV-RNA in den Huh-7-Zellen zu bestimmen, wurde die Gesamt-RNA isoliert und mittels des gängigen Northern-Blot Verfahrens unter Verwendung einer Plusstrang-spezifischen Ribosonde (= RNA-Sonde) analysiert. Hierfür wurde von den jeweiligen Zellklonen Gesamt-RNA nach der Methode von Chomczynski und Sacchi 1987, *Anal. Biochem.* 162, 156 isoliert, und 10 µg RNA, was dem Gesamt-RNA-Gehalt von 0,5 - 1 x 10⁵ Zellen entspricht, mittels denaturierender Formaldehyd-Agarose-Gelelektrophorese aufgetrennt (Spuren 3 bis 12 der Fig. 1 B). Als Größenmarker mit authentischer Sequenz wurden gleichzeitig 10³ in-vitro-Transkripte (vtr.), die zu den I₃₅₉/NS2-3'/wt oder den I₃₅₉/NS3-3'/wt Replikon-RNAs korrespondieren, mit aufgetrennt (Spur 1 bzw. Spur 2). Die aufgetrennte RNA wurde auf Nylon-Membranen transferiert und mit radioaktiv markierter Plusstrang-spezifischer RNA-Sonde, die komplementär zu dem kompletten NPT-Gen und der HCV-IRES von Nukleotid 377 bis Nukleotid 1 war, hybridisiert. Die Positionen der HCV-spezifischen RNAs (Pfeile) und der 28S rRNA sind rechts von Spur 12 angegeben, die Größen (Anzahlen der Nukleotide) der RNA-Marker sind links von Spur 1 angegeben. Die RNA Marker-Fragmente enthalten HCV-Sequenzen und hybridisieren deshalb mit der Ribosonde (= RNA-Sonde). Die Ergebnisse dieser Analyse sind in Fig. 1 B dargestellt.

[0068] Mit Ausnahme des mit dem defekten HCV-Genom-Konstrukt transfizierten Klons 8-1, lieferten alle Zellklone homogene HCV-RNAs korrekter Länge (ca. 8640 Nukleotide im Fall des NS2-3' und ca. 7970 Nukleotide im Fall des NS3-3' Replikons). Dieser Befund ist ein Indiz dafür, daß die funktionalen Replikons bzw. die funktionalen HCV-Genom-Konstrukte die G418 Resistenz übertragen. Um auszuschließen, daß die G418 Resistenz auf eine Plasmid-DNA zurückzuführen ist, die in das Genom der Huh-7 Wirtszelle integriert ist und unter der Kontrolle eines zellulären Promotors transkribiert wird, wurde von jedem Klon die DNA mittels einer NPT-Gen-spezifischen PCR untersucht. Hierbei wurde aus den selektierten Huh-7-Zellklonen die DNA mittels Verdau mit Proteinase K (40 µg/ml, 1h, 37°C) in 10 mM Tris, pH 7,5, 1 mM EDTA, 0,5% SDS und anschließender Extraktion mit Phenol, Phenol/Chloroform und Isopropanolpräzipitation isoliert. Das DNA-Präzipitat wurde in 10 mM Tris (pH 7,5) und 1 mM EDTA gelöst und 1 Stunde mit RNase A inkubiert. Im Anschluß an eine Phenol/Chloroform Extraktion und Ethanol Präzipitation wurde 1 µg DNA, entsprechend 4 - 8 x 10⁴ Zellen, mittels PCR unter Einsatz NPT-Gen-spezifischer Primer (5'-TCAAGACCGACCTG TCGGGTGGCC-3' und 5'-CTTGAGCCCTGGGGAACAGTTCTGGC-3') analysiert und ein DNA-Fragment bestehend aus 379 Nukleotiden erzeugt. Die Spezifität des PCR-Produkts wurde mittels Southern Blot Verfahren nachgewiesen, wobei ein Digoxigenin-markiertes DNA Fragment eingesetzt wurde, das zu dem NPT-Gen korrespondiert. Als Positiv-Kontrollen (zum Nachweis etwa vorhandener kontaminierender Nukleinsäuren) wurde das PCR-Verfahren mit 10⁷ Plasmid Molekülen oder 1 µg DNA aus einer BHK Zelllinie, die stabil mit einem Neomycin-Resistenz-Gen transfiziert war, durchgeführt, und als Negativ-Kontrolle wurde die PCR mit denselben Reagenzien aber ohne zugesetzte DNA durchgeführt.

Die Ergebnisse dieser Untersuchung sind in Fig. 1 C dargestellt. Die Spuren 1 und 2 repräsentieren die Positiv-Kontroll-

len, Spur 13 repräsentiert die Negativ-Kontrolle. Die Zahlenangaben links der Spur 1 bezeichnen die Größe der Nukleotid-Marker-Moleküle. Außer in Klon 7-3 (Fig. 1C, Spur 3), der von Zellen nach Transfektion mit einem NS2-3' Replikon/NS2-3'HCV-Genom-Konstrukt stammt, und in Klon 8-1 (Fig. 1C, Spur 12), der von Zellen nach Transfektion mit einem defekten HCV-Genom-Konstrukt stammt, war in keinem Zellklon eine NPT-DNA nachweisbar. Dieser Befund ist ein weiteres Indiz dafür, daß die G418 Resistenz der meisten Klone durch die replizierende HCV-RNA vermittelt wurde. Aber auch unabhängig von diesen Ergebnissen ist es unwahrscheinlich, daß HCV-RNAs mit korrekter Größe von integrierter Plasmid DNA erzeugt wird, denn die für die in-vitro-Transkription verwendeten Plasmide enthalten weder einen eukaryotischen Promotor noch ein Polyadenylierungssignal. Im Fall des Klons 7-3 ist die Resistenz deshalb höchst wahrscheinlich sowohl durch das HCV-RNA-Konstrukt bzw. die replizierende HCV-RNA als auch durch eine integrierte NPT DNA Sequenz vermittelt worden, während die Resistenz der Zellen von Klon 8-1 ausschließlich auf die integrierte Plasmid DNA zurückzuführen ist.

[0069] Um zu bestätigen, daß die G418 Resistenz von einer autonom replizierenden HCV-RNA vermittelt ist, wurde der Klon 9-13 (Fig. 1 B, Spur 11) weiteren Tests unterworfen. Klon 8-1, der integrierte Kopien des NPT-Gens trägt, wurde überall als Negativkontrolle eingesetzt. Mit dem Ziel, die Anwesenheit von NPT-DNA im Klon 9-13 rigoros auszuschließen, wurde eine PCR durchgeführt, die den Nachweis von < 1000 NPT-Gen-Kopien in ~ 40.000 Zellen erlaubt. Das Ergebnis dieser PCR ist in Fig. 2A dargestellt. Im einzelnen wurde bei dieser PCR wie folgt verfahren: Es wurden jeweils $10^5 - 10^6$ Plasmid Moleküle ($1_{377}/NS3-3'/wt$) entweder direkt (Spuren 7 - 11) oder nach Zugabe von jeweils $1 \mu g$ 9-13 DNA (Spuren 2 - 6) in dem Test eingesetzt. Die Spezifität der amplifizierten DNA Fragment wurde mittels Southern Blot unter Verwendung einer NPT-spezifischen Sonde bestimmt. Eine PCR ohne DNA-Sonde wurde als Negativ-Kontrolle durchgeführt (Spur 12).

Selbst mit dieser sensitiven Methode wurde in einem μg DNA des Zellklons 9-13 keine Plasmid DNA gefunden (Spur 1). Um die Menge an HCV Plus- und Minusstrang RNAs in diesen Zellen abzuschätzen, wurde eine Verdünnungsreihe von Gesamt-RNA mit dem Northern-Blot-Verfahren unter Verwendung einer Plus- oder Minusstrang-spezifischen radioaktiv markierten Ribosonde (= RNA-Sonde) analysiert. Hierfür wurden jeweils 8, 4 oder 2 μg Gesamt-RNA, die aus den Zellklonen 9-13 und 8-1 isoliert worden waren, parallel zu bekannten Mengen analoger in-vitro-Transkripte mit Plus- oder Minusstrang-Polarität (Kontroll-RNAs) im Northern-Blot-Verfahren analysiert und anschließend einer Hybridisierung unterworfen. Die Hybridisierung wurde mit einer Plusstrang-spezifischen Ribosonde, die das komplette NPT-Gen und die HCV-IRES abdeckte ('plusstrang', obere Bildtafel), oder mit einer Minusstrang-spezifischen RNA-Sonde, die zu der NS3-Sequenz komplementär war ('minusstrang', untere Bildtafel) durchgeführt. Die Pfeile markieren die Positionen von Replikon-RNA. Die Ergebnisse dieser Analyse sind in Fig. 2 B dargestellt.

Im Fall des Plusstrangs wurden ca. 10^5 Kopien/ μg Gesamt-RNA nachgewiesen, was 1000 - 5000 HCV-RNA-Molekülen pro Zelle entspricht, während die Menge an Minusstrang-RNA 5- bis 10-fach niedriger war. Dieses Ergebnis stimmt mit der Annahme überein, daß die Minusstrang RNA die replikative Zwischenform bzw. Zwischenkopie ist, die als Vorlage für die Synthese der Plusstrang Moleküle dient.

Da die Reaktion im wesentlichen von der viralen RNA-abhängigen RNA Polymerase katalysiert wird, sollte die Synthese der HCV-RNAs resistent gegen Dactinomycin sein, einem Antibiotikum, das selektiv die RNA-Synthese von DNA-Matrizen inhibiert, nicht jedoch die RNA-Synthese von RNA-Matrizen. Um diese Vermutung zu bestätigen, wurden Zellen mit [3H] Uridin in Anwesenheit von Dactinomycin inkubiert, die radioaktiv markierten RNAs extrahiert, mittels denaturierender Agarose-Gel-Elektrophorese aufgetrennt und mit Hilfe eines handelsüblichen Bio-Imagers unter Verwendung einer [3H]-sensitiven Bildplatte analysiert. Hierfür wurden jeweils ca. 5×10^5 Zellen der Klone 9-13 und 8-1 mit $100 \mu Ci$ [3H]Uridin für 16 Std. in Abwesenheit (-) oder Gegenwart (+) von $4 \mu g/ml$ Dactinomycin (Dact) inkubiert. Im Anschluß an diese Markierungsreaktion wurde die Gesamt-RNA präpariert und mittels Formaldehyd-Agarose-Gel-Elektrophorese analysiert. In den beiden ersten Spuren ist nur 1/10 der Gesamt-RNA dargestellt. Die radioaktiv markierte RNA wurden mit einem BAS-2500 Bio-Imager (Firma Fuji) sichtbar gemacht.

Die Ergebnisse dieser Analyse sind in Fig. 2 C dargestellt. In Übereinstimmung mit dem Inhibitor-Profil der NSSB Polymerase (Behrens et al., 1996, *EMBOJ.* 15, 12 und Lohmann et al., 1997, *J Virol.* 71, 8416) war die Replikation der HCV RNA nicht durch Dactinomycin beeinflusst worden, während die Synthese von zellulärer RNA gehemmt worden war. Um die Identität der viralen RNA zu bestätigen, wurde eine RT-PCR zur Reklonierung der replizierten Sequenzen durchgeführt. Die Sequenzanalyse der reklonierten RNA zeigte, daß die RNA in dem Klon 9-13 HCV-spezifisch ist und mit dem transkribierten Transkript des HCV-Konstrukts $1_{377}/NS3-3'/wt$ übereinstimmt.

[0070] Zur Analyse der viralen Proteine wurden die betreffenden Zellen zunächst metabolisch mit [^{35}S] Methionin/Cystein radioaktiv markiert, anschließend lysiert und danach die HCV-spezifischen Proteine mittels Immunopräzipitation aus dem Zell-Lysaten isoliert. Die Ergebnisse dieser Analysen sind in Fig. 3 A dargestellt. Im einzelnen wurde dabei wie folgt verfahren: Zellen der Zellklone 9-13 (wt) und 8-1 (Δ) waren durch Behandlung für 16 Stunden mit einer dem Fachmann geläufigen und im Handel erhältlichen Protein-Markierungs-Mischung (z.B. NEN Life Science) metabolisch radioaktiv markiert worden. Mittels Immunopräzipitation (IP) unter nicht-denaturierenden Bedingungen (z.B. nach Bartschlagel et al., 1995, *J. Virol.* 69, 7519) und unter Verwendung von drei verschiedenen Antisera (3/4, 5A, 5B, gemäß Markierung am oberen Ende der Spuren 1 bis 12) waren die HCV-spezifischen Proteine vom Zell-Lysat abge-

trennt worden. Die Immunokomplexe wurden mittels Tricine SDS-PAGE analysiert und mittels Autoradiographie sichtbar gemacht. Um authentische Größenmarker zu erhalten, wurde das homologe Replikonkonstrukt $\text{I}_{777}/\text{NS3}/\text{wt}$ einer transienten Expression mit dem Vaccinia Virus T7-Hybrid System in Huh-7 Zellen unterworfen. Die dabei erhaltenen Produkte waren als Größenmarker (Spuren 7 - 9) parallel zu den Zellen der Klone 9-13 und 8-1 behandelt worden. Identifizierte HCV-Proteine sind am linken Rand von Spur 1 markiert, die Molekulargewichte (in Kilodalton) sind am rechten Rand von Spur 9 angegeben. Es ist anzumerken, daß das verwendete NS3/4-spezifische Antiserum (3/4) bevorzugt mit NS4A und NS4B reagiert, was zu einer Unterrepräsentation von NS3 führt.

[0071] Alle viralen Antigene waren eindeutig nachweisbar und ihre apparenten Molekulargewichte zeigten keine Abweichungen gegenüber denjenigen, die nach transientser Expression desselben bicistronischen HCV-RNA-Konstrukts in den ursprünglichen Huh-7 Zellen ermittelt wurden. Um die subzelluläre Verteilung der viralen Antigene zu bestimmen, wurde eine Immunofluoreszenz-Nachweisreaktion unter Einsatz von NS3- und NS5A-spezifischen Antisera durchgeführt (z.B. nach Bartschlagler et al., 1995, *J. Virol.* 69, 7519). Hierfür wurden Zellen der Klone 9-13 (wt) und 8-1 (Δ) 24 Std. nach dem Aussäen auf Deckgläsern mit Methanol/Azetan fixiert und mit polyklonalen NS3- oder NS5A-spezifischen Antisera inkubiert. Die gebundenen Antikörper wurden mit einem kommerziell erhältlichen FITC-konjugierten Anti-Kaninchen-Antiserum sichtbar gemacht. Zur Unterdrückung unspezifischer Fluoreszenzsignale wurden die Zellen mit dem Farbstoff 'Evans Blue' gegengefärbt.

[0072] Die Ergebnisse dieses Nachweistests sind in Fig. 3 B dargestellt. Mit beiden Antisera war eine starke Fluoreszenz im Zytoplasma nachweisbar. Die NS5A-spezifischen Antisera führten außerdem zu einer schwachen Zellkern-Fluoreszenz, was darauf hindeutet, daß zumindest kleine Mengen dieses Antigens auch zum Zellkern gelangen. Die generell dominierende Präsenz der viralen Antigene im Zytoplasma ist jedoch ein starkes Indiz dafür, daß die HCV-RNA Replikation im Zytoplasma stattfindet — so wie das bei den meisten RNA-Viren der Fall ist.

[0073] Diese Ergebnisse belegen klar, daß mit dem hier beschriebenen Versuchsansatz der Aufbau eines Zellkultursystems für das HCV gelungen ist, dessen Effizienz alles bisher bekannte um Größenordnungen übersteigt und erstmalig den Nachweis viraler Nukleinsäuren und Proteine mit konventionellen und bewährten biochemischen Methoden erlaubt. Erst diese Effizienz erlaubt überhaupt detaillierte Untersuchungen der HCV-Pathogenese, genetische Analysen verschiedener HCV-Funktionen und ein genaues Studium der Virus-/Wirtszellwechselwirkungen, wodurch sich neue Ansatzpunkte für die Entwicklung einer antiviralen Therapie definieren lassen.

Beispiel 3: Transfektion von Huh-7 Zellen mit HCV-Genomkonstrukten

[0074] Huh-7 Zellen werden wie in Beispiel 2 beschrieben transifiziert und selektioniert, wobei hier jedoch selektionierbare Konstrukte verwendet werden, die das vollständige Virusgenom enthalten. Die erhaltenen Zellklone werden analog dem Beispiel 2 mittels PCR auf Abwesenheit von HCV-DNA untersucht und die produktive Replikation der HCV-RNA wird danach mittels Northern Blot, [^3H]Uridinmarkierung in Abwesenheit von Dactinomycin, Nachweis der viralen Proteine bzw. Antigene vorzugsweise mit Hilfe des Western Blots, der Immunopräzipitation oder der Immunfluoreszenz nachgewiesen. Im Gegensatz zu den im Beispiel 2 beschriebenen Ansätzen lassen sich mit dem hier beschriebenen Konstrukt außerdem vollständige und sehr wahrscheinlich infektiöse Viren erhalten, was bei den dort (in Beispiel 2) beschriebenen Subgenomkonstrukten nicht der Fall ist. Diese Viren, die in der Zelle und dem Zellkulturüberstand vorhanden sind, werden beispielsweise mittels Ultrazentrifugation, Immunopräzipitation oder Fällung mit Polyethylenglykol konzentriert und alle exogenen, d.h. nicht im Viruspartikel eingebauten Nukleinsäuren werden mittels Inkubation mit Nukleasen (RNase, DNase, Mikrococcusnuklease) verdaut. Auf diese Weise lassen sich alle kontaminierenden Nukleinsäuren, die nicht im schützenden Viruspartikel enthalten sind, entfernen. Die geschützte virale RNA wird nach Inaktivierung der Nukleasen, beispielsweise mittels Inkubation mit Proteinase K in einem SDS-haltigen Puffer durch Extraktion mit Phenol und Phenol/Chloroform isoliert und mittels Northern Blot oder RT-PCR unter Verwendung HCV-spezifischer Primer nachgewiesen. Auch in diesem Versuchsansatz ist die Kombination des beschriebenen HCV-Konsensusgenoms mit einem Selektionsmarker entscheidend für die effiziente Produktion von viraler RNA, viralem Protein und damit von HCV-Partikeln.

Beispiel 4: Herstellung und Anwendung eines HCV-RNA Konstrukts, bei dem das Resistenzgen über ein Ribozym bzw. eine Erkennungsstelle für ein Ribozym mit der HCV-Subgenom-Sequenz verbunden ist.

[0075] Es wird ein HCV-RNA-Konstrukt gemäß Beispiel 1 oder Beispiel 3 hergestellt, bei dem ein Antibiotikumresistenzgen über ein Ribozym bzw. eine Erkennungsstelle für ein Ribozym mit der HCV-RNA-Sequenz verbunden ist. Solche Konstrukte sind in Fig. 7 schematisch dargestellt. Huh-7 Zellen werden wie in Beispiel 2 beschrieben mit diesem HCV-RNA-Konstrukt transifiziert. Nach der Transfektion in die Zellen erfolgt zunächst die Selektion mit dem entsprechenden Antibiotikum. In den dabei erhaltenen Zellklonen wird das einklonierte Ribozym aktiviert oder, im Fall eines Konstrukts, das eine Erkennungsstelle für ein Ribozym trägt, wird das Ribozym in die Zelle eingeschleust (z.B. mittels Transfektion eines Ribozymkonstrukts oder Infektion mit einem viralen Expressionsvektor, in den das entsprechende

Ribozym eingesetzt wurde). In beiden Fällen wird durch die ribozymvermittelte Spaltung das Resistenzgen von der HCV-RNA-Sequenz abgetrennt. Das Ergebnis ist im Fall des HCV-Genom-Konstrukts ein authentisches HCV-Genom ohne Resistenzgen, das zur Bildung authentischer infektiöser Viruspartikel befähigt ist. Im Fall des HCV-Subgenom-Konstrukts entsteht ein HCV-Replikon ohne Resistenzgen.

Beispiel 5: Kointransfektion eines HCV-RNA-Konstrukts mit einem separaten Luciferase-Transfektionskonstrukt

[0076] Es wird ein HCV-RNA-Konstrukt gemäß Beispiel 1 (A) oder Beispiel 3 oder Beispiel 4 hergestellt. Parallel dazu wird ein Transfektionskonstrukt hergestellt, welches das Luciferasegen umfaßt, wobei dieses Luciferasegen vermittels einer ersten Nukleotidsequenz, die für eine HCV-Protease- (z.B. NS3-Protease-) Spaltungsstelle kodiert, mit einer zweiten Nukleotidsequenz, die für ein anderes Protein oder einen Teil eines anderen Proteins kodiert, verbunden ist. HCV-RNA-Konstrukt und Transfektionskonstrukt werden in beliebige Wirtszellen, vorzugsweise Hepatomazellen, insbesondere Huh-7-Zellen, eingeschleust. Dies kann auf die in Beispiel 2 beschriebene Art und Weise geschehen. Das Produkt des modifizierten Luciferasegens ist ein Luciferase-Fusionsprotein, in dem die Luciferase auf Grund der Fusion mit dem Fremdanteil inaktiv ist. In transfizierten Zellen mit hoher HCV-Replikation wird das Fusionsprotein, das ja eine Schnittstelle für eine HCV-Protease enthält, gespalten und damit die aktive Form der Luciferase freigesetzt, die sich durch luminometrische Messung bestimmen läßt. Wird die Replikation des HCV-RNA-Konstrukts gehemmt, wird das Fusionsprotein nicht gespalten und keine aktive Luciferase freigesetzt. Infolgedessen ist die quantitative Bestimmung der Luciferase ein Maß für die Replikation des HCV-Subgenom-Konstrukts. Anstelle des Luciferasegens kann ebenso ein anderes Reportergen verwendet werden, das in analoger Weise modifiziert ist, so daß seine Expression von der Virusreplikation abhängt, obwohl dieses Reportergen nicht Bestandteil des HCV-Subgenom-Konstrukts ist. Es kann auch ein zelluläres Protein, welches durch die HCV-Proteine oder Nukleinsäure inaktiviert oder aktiviert wird, als sogenannter Surrogatmarker verwendet werden. In diesem Fall ist die Expression bzw. Aktivität dieses Surrogatmarkers ein Maß für die Replikation der viralen DNA.

Beispiel 6: Herstellung von HCV-Subgenom-Konstrukten mit integrierten Fremdgenen zur Verwendung als Leberzell-spezifische Gentherapien für die Gentherapie

[0077] Diese rekombinanten und selektionierbaren HCV-Subgenom-Konstrukte werden in trans-komplementierende Helferzelllinien transfiziert, d.h. in Zelllinien, die induzierbar oder konstitutiv die fehlenden Funktionen (beispielsweise die Strukturproteine) exprimieren. Zellklone, die ein funktionelles HCV-Subgenom-Konstrukt enthalten, lassen sich durch entsprechende Selektion etablieren. Die von der Wirtszelle exprimierten Virus-Strukturproteine erlauben die Bildung von Viruspartikeln, in die die RNA der HCV-Subgenom-Konstrukte eingeschleust wird. Das Ergebnis sind also virus-ähnliche Partikel, die ein erfindungsgemäßes HCV-Subgenom-Konstrukt einschließlich des inkorporierten Fremdgens enthalten und die dieses mittels Infektion auf andere Zellen übertragen können. Ein Beispiel für ein solches Konstrukt ist in Fig. 8 dargestellt. Es besteht auch die Möglichkeit, das hier beschriebene erfindungsgemäße HCV-Subgenom-Konstrukt mit integriertem Fremdgen direkt als Expressionsvektor einzusetzen. Dabei wird analog dem vorgenannten Verfahren vorgegangen, allerdings mit dem Unterschied, daß Zelllinien transfiziert werden, die keine trans-komplementierenden Faktoren exprimieren. In diesem Fall dient das HCV-Konstrukt also lediglich als Expressionsvektor.

Beispiel 7: Herstellung zellkultur-adaptierter HCV-RNA-Konstrukte

(A) Isolationsverfahren

[0078] Für die Bestimmung adaptiver Mutationen und die Herstellung zellkultur-adaptierter HCV-RNA-Konstrukte wurde wie folgt verfahren: Zellen wurden mit einem HCV-RNA-Konstrukt wie unter den Beispielen 1 und 2 beschrieben transfiziert und G418-resistente Zellklone hergestellt. Zur Bestimmung der Replikationskompetenz (darunter wird in diesem Zusammenhang die Anzahl G418-resistenter Zellklone verstanden, die pro Mikrogramm transfizierter HCV-RNA bzw. HCV-RNA-Konstrukt erhalten wird) wurde exemplarisch die Gesamt-RNA aus einem der Zellklone, genannt 9-13 (Fig. 1B, Spur 11) isoliert und die Menge der darin enthaltenen HCV-RNA mittels Northern-blot wie in Fig. 2 B beschrieben bestimmt. 10 Mikrogramm der Gesamt-RNA, die ca. 10^9 Moleküle HCV-RNA enthielt, wurde anschließend per Elektroporation in naive Huh-7 Zellen eingeschleust (Fig. 9). Parallel dazu wurden 10^9 in vitro Transkripte der analogen neo-HCV-RNA, die mit isolierter Gesamt-RNA aus naiven Huh-7 Zellen auf eine Gesamt-RNA-Menge von 10 µg aufgefällt worden war, in naive Huh-7 Zellen transfiziert. Nach Selektion mit G418 wurde die Anzahl der Zellkolonien, ausgedrückt in "colony forming units (cfu) pro Mikrogramm RNA" in den beiden Ansätzen bestimmt. Bei einer Konzentration von 500 µg/ml G418 im Selektionsmedium betrug die Zahl der Kolonien, die mit der in der isolierten Gesamt-RNA enthaltenen HCV-RNA erhalten wurde, ca. 100.000 cfu pro Mikrogramm HCV-RNA. Dagegen wurden mit dersel-

ben Menge in vitro transkribierter HCV-RNA nur 30 - 50 Kolonien erhalten. Dieses Ergebnis belegt, daß die spezifische Infektiosität der HCV-RNA, die aus den Zellklonen isoliert wurde, ca. 1.000 - 10.000-fach höher ist als die Infektiosität der analogen in vitro Transkripte. Das methodische Vorgehen ist in Fig. 9 dargestellt.

[0079] Mit Hilfe der 'long-distance RT-PCR' wurde die HCV-RNA aus der Gesamt-RNA der 9-13 Zellen amplifiziert, die PCR-Amplifikate kloniert und zahlreiche Klone sequenziert. Ein Vergleich der Sequenzen dieser reklonierten RNAs mit der Sequenz der RNA, die ursprünglich in die naiven Huh-7 Zellen eingeschleust wurde ergab, daß die reklonierten RNAs zahlreiche Aminosäureaustausche besaßen, die über die gesamte HCV-Sequenz verteilt waren (Fig. 10). *Sil*-Fragmente dieser reklonierten Mutanten wurden im Austausch gegen das analoge *Sil*-Fragment des ursprünglichen Replikonkonstrukts in dieses eingeführt und RNAs der jeweiligen Mutanten wurden in naive Huh-7 Zellen eingeschleust. Nach Selektion mit G418 wurde dann für jede HCV-RNA-Mutante die Zahl der gebildeten Kolonien bestimmt. Während mit der Ausgangs-RNA nur 30 - 50 Kolonien pro Mikrogramm RNA erhalten wurde war die Koloniezahl bei zwei der reklonierten Varianten deutlich höher (Fig. 10). Im Fall der HCV-RNA-Konstrukte 9-13I und 9-13C betrug die spezifische Infektiosität 100 - 1.000 cfu pro Mikrogramm RNA und beim 9-13F Replikon sogar 1.000 - 10.000 cfu pro Mikrogramm RNA. Diese Ergebnisse zeigen, daß die Aminosäureaustausche in dem analysierten NS3-5B-Bereich der Mutanten 9-13I, 9-13C und insbesondere 9-13F zu einer deutlichen Erhöhung der Replikationskompetenz führten. Demgegenüber waren alle anderen HCV-RNA-Konstrukte (9-13 A, B, G, H und K) nicht mehr replikationskompetent, enthielten also letale Mutationen.

[0080] Zwecks Beantwortung der Frage, welche der Aminosäureaustausche im 9-13F-Konstrukt zur Steigerung der Replikation führten, wurden die Austausche einzeln oder in Kombination in das Ausgangs-HCV-RNA-Konstrukt eingeführt und die entsprechenden RNAs in naive Huh-7 Zellen eingeschleust. Das Ergebnis der Transfektionen mit diesen RNAs ist in Tabelle 1 zusammengefaßt. Daraus wird ersichtlich, daß im vorliegenden Beispiel die hohe Replikationskompetenz durch mehrere Mutationen bedingt ist. Den größten Beitrag leisten die Aminosäureaustausche in den HCV-RNA-Abschnitten NS5A und NS4B. Auch die einzelnen Austausche in der NS3-Region leisten einen Beitrag, der möglicherweise auf einem Synergismus dieser Einzelaustausche beruht.

Diese Befunde belegen, daß es durch die G418-Selektion der Zellen, die mit den neo-HCV-RNA-Konstrukten transfiziert wurden, zur Anreicherung solcher HCV-RNAs kam, die eine deutlich höhere Replikationskompetenz hatten. Mit dem hier beschriebenen Versuchsansatz lassen sich HCV-RNA-Konstrukte mit sehr unterschiedlicher Replikationseffizienz selektionieren. Je höher die Konzentration des Antibiotikums in dem Selektionsmedium ist, in/auf dem die HCV-RNA-Konstrukt-haltigen Zellen zwecks Selektion kultiviert werden, desto höher muß der Grad an adaptiven Mutationen und damit die Replikationseffizienz in den betreffenden HCV-RNA-Konstrukten sein, damit die Zellen auswachsen können. Werden die Selektionen mit niedrigeren Antibiotikum-Konzentrationen durchgeführt, können auch solche Zellen überleben und sich vermehren, die im Vergleich geringer adaptive Mutationen und eine weniger hohe Replikationseffizienz aufweisen.

Das bisher beschriebene HCV-RNA-Konstrukt 9-13F, das mehrere adaptive Mutationen enthielt, hatte eine erwiesenermaßen höhere Replikationseffizienz als die parentale HCV-RNA. Um HCV-RNAs mit noch höherer Replikation in Zellkultur zu erhalten, wurde die HCV-RNA, die in der Gesamt-RNA eines ausgewählten Zellklons enthalten war, mehrfach in naiven Huh-7 Zellen passagiert. Dieser ausgewählte Zellklon, genannt 5-15, wurde durch Transfektion mit dem HCV-RNA-Konstrukt $\text{I}_{\text{NS3-3'}}$ erhalten (Fig. 1). Er entspricht weitgehend dem Zellklon 9-13, der durch Transfektion mit einem HCV-RNA-Konstrukt hergestellt wurde, das eine um 22 Nukleotide kürzere HCV-IREs besaß ($\text{I}_{\text{NS7-NS3-3'}}$; Fig. 1). 10 Mikrogramm Gesamt-RNA, isoliert aus dem Zellklon 5-15, wurden mittels Elektroporation in naive Huh-7 Zellen eingeschleust und die Zellen einer Selektion mit 1 mg/ml G418 unterzogen. Aus einem der so erzeugten Zellklone wurde wiederum Gesamt-RNA isoliert, in naive Huh-7 Zellen transfiziert und analog selektioniert. Dieser Vorgang wurde insgesamt viermal wiederholt. Nach der vierten Passage wurde aus einem Zellklon die Gesamt-RNA isoliert und die neo-HCV-RNA mit Hilfe der 'long-distance RT-PCR' amplifiziert. Das amplifizierte DNA-Fragment wurde mit dem Restriktionsenzym *Sfi* verdaut und in das *Sfi*-restriktierte Ausgangskonstrukt $\text{I}_{\text{NS3-3'}}$ inseriert. Insgesamt wurden über 100 DNA-Klone erhalten und zunächst mittels Restriktionsverdau analysiert. In vitro transkribierte RNA von ca. 80 dieser Klone wurde jeweils in naive Huh-7 eingeschleust und einer Selektion mit 500mg/ml G418 unterzogen. Von den 80 untersuchten neo-HCV-RNA-Varianten erwiesen sich die allermeisten als replikationsdefekt. Bei zwei Mutanten, genannt 5.1 und 19, war die spezifische Infektiosität, ausgedrückt als 'colony forming units' pro Mikrogramm RNA, jedoch sehr deutlich erhöht (Tabelle 2). Durch mehrfache Passage der RNA in Zellkultur lassen sich offensichtlich HCV-RNAs herstellen, deren Replikationseffizienz aufgrund von Mutationen (sog. 'adaptiven Mutationen') mehrere Größenordnungen höher ist als die der ursprünglich aus dem Patienten klonierten RNAs.

(B) Modifikationsverfahren

[0081] Solche nach (A) erzeugten und identifizierten adaptiven Mutationen können in ein wenig replikationskompetentes HCV-RNA-Konstrukt übertragen werden und führen zu einer massiven Steigerung der Replikation dieses Konstrukts. Diese Steigerung ist so hoch, daß damit nachweislich HCV-RNAs in Zellkultur zur Replikation gebracht

werden können, die kein selektierbares Markergen mehr besitzen. Fig. 12 zeigt einen Vergleich der Replikationseffizienz von HCV-RNAs, die entweder der Ausgangssequenz oder den adaptierten Sequenzen 9-13F bzw. 5.1 entsprechen. Zwecks einfacher Messung wurde das *neo*-Gen entfernt und durch das Gen für die Luziferase ersetzt. Als Negativkontrolle diente wiederum ein HCV-RNA-Konstrukt, das auf Grund einer inaktivierenden Mutation der NSSB RNA-Polymerase replikationsdefekt war. Schon 24 Stunden nach der Transfektion erkennt man einen deutlichen Unterschied in der Luziferaseaktivität zwischen der defekten RNA und den 9-13F bzw. 5.1-Konstrukten während zwischen der defekten RNA (318 DN) und dem Ausgangs-RNA-Konstrukt (wt) das keine adaptiven Mutationen besaß, kaum ein Unterschied zu sehen war. Während des gesamten Beobachtungszeitraums wurde die höchste Luziferaseaktivität und damit die höchste Replikation mit der 5.1-RNA erhalten. Diese Befunde belegen nicht nur die hohe Replikationseffizienz dieser RNA, sondern zeigen auch, daß es möglich ist, mit adaptierten HCV-RNA-Konstrukten ein Zellkultursystem aufzubauen, für das die Anwesenheit eines selektierbaren Gens nicht mehr notwendig ist. Eine zusammenfassende Übersicht der Nukleotid- und Aminosäureunterschiede zwischen dem Ausgangskonstrukt und den Mutanten 9-13F, 5.1 und 19 ist in Tabelle 3 gegeben.

15 Beispiel 8: Herstellung zellkultur-adaptierter HCV-RNA-Vollängengenome

[0082] In den Beispielen 1 bis 7 wurde stets eine subgenomische HCV-RNA verwendet, der die gesamte Strukturproteinregion von Core bis einschließlich p7 bzw. NS2 fehlte. Im vorliegenden Beispiel 8 wird gezeigt, daß es möglich ist, mit Hilfe der adaptierten NS3-5B-Sequenz ein HCV-Vollängengenom in Zellkultur zur Replikation zu bringen. Zu diesem Zweck wurde zunächst das SfiI-Fragment der gemäß Beispiel 7 hergestellten, hoch adaptierten HCV-RNA 5.1 in ein selektionierbares HCV-Vollängengenom transfiziert (Fig. 12). Dieses HCV-Genom wurde in naive Huh-7 Zellen transfiziert und einer Selektion mit unterschiedlichen G418-Konzentrationen unterzogen. In Abhängigkeit von der Selektionsstärke (der G418-Konzentration) wurde eine unterschiedlich große Zahl an Zellklonen erhalten (Fig. 12 B). Im Vergleich dazu wurden mit dem unveränderten HCV-Vollängengenom, das keine adaptiven Mutationen enthielt, keine Kolonien erhalten, ebenso mit der Negativkontrolle, die auf Grund einer inaktivierenden Mutation in der NSSB RNA-Polymerase replikationsdefekt war. Zum Nachweis dafür, daß die so entstandenen Zellklone tatsächlich ein autonom replizierendes HCV-Vollängengenom enthielten, wurde Gesamt-RNA aus mehreren Zellklonen isoliert und mittels Northern-Blot analysiert. In allen Zellklonen war die Vollängen HCV-RNA eindeutig nachweisbar (Fig. 12). Damit ist eindeutig belegt, daß es mit Hilfe der an Zellkulturen adaptierten HCV-Sequenzen möglich ist, ein HCV-Vollängengenom herzustellen, das mit hoher Effizienz und autonom in einer Zelllinie repliziert, d.h. es können mit dem erfindungsgemäßen System auch adaptierte HCV-Vollängengenome hergestellt werden. Da dieser Klon darüber hinaus die vollständige HCV-Sequenz besitzt, also auch die für die Viruspartikelbildung notwendigen Strukturproteine, ist es mit diesem System möglich, große Mengen infektiöser Viruspartikel in Zellkulturen herzustellen. Zum Nachweis dieser Viren werden zellfreie Überstände von Zellen, die ein replizierendes HCV-Vollängengenom tragen, auf naive Huh-7 Zellen gegeben und die so infizierten Zellen einer Selektion mit G418 unterzogen. Jeder Zellklon, der unter diesen Bedingungen auswächst, geht auf eine infizierte Zelle zurück. Die Viren in den Zellkulturüberständen von Zellen, die ein replizierendes HCV-Vollängengenom besitzen, können aber auch mit verschiedenen im Stand der Technik bekannten Verfahren wie Ultrazentrifugation oder Mikrodialyse angereichert und gereinigt werden und dann zur Infektion naiver Zellen verwendet werden. Mit diesem Verfahren ist eindeutig gezeigt, daß mit dem erfindungsgemäßen HCV-Zellkultursystem zellkultur-adaptierte HCV-Vollängengenome hergestellt werden können, die mit hoher Effizienz in Zellen replizieren und infektiöse Viren produzieren. Diese können ebenfalls durch Infektion eines Versuchstiers, vorzugsweise dem Schimpansen, nachgewiesen werden.

45 Beispiel 9: Herstellung von HCV-Vollängen-Konstrukten und HCV-Subgenom-Konstrukten mit Reportergen.

[0083] Es wird ein HCV-RNA-Konstrukt hergestellt, bei dem anstelle des Antibiotikumresistenzgens ein Reportergen eingefügt wird (Fig. 13). Dabei kann die Replikation anhand der Menge bzw. der Aktivität des Reportergens bzw. Reportergenprodukts bestimmt werden. Das Reportergen ist vorzugsweise ein Gen aus der Gruppe der Luziferasegene, dem CAT-Gen (Chloramphenicol-Acetyl-Transferase-Gen), dem lacZ-Gen (beta-Galaktosidasegen), dem GFP-Gen (green fluorescence protein Gen), dem GUS-Gen (Glukuronidasegen) oder dem SEAP-Gen (sezernierte alkalische Phosphatasegen). Diese Reportergene bzw. deren Produkte, nämlich die entsprechenden Reporterproteine, können z.B. mittels Fluoreszenz, Chemilumineszenz, colorimetrisch oder mit Hilfe immunologischer Methoden (z.B. enzyme-linked immunosorbent assay, ELISA) bestimmt werden. Das Reportergen kann entweder von einer eigenen IRES exprimiert werden oder in Form eines Fusionsproteins, das entweder als solches aktiv ist oder mittels einer proteolytisch spaltbaren Aminosäuresequenz so mit einem HCV-Protein verbunden ist, daß es von einer zellulären oder viralen (HCV-)Protease von diesem abgespalten wird.

Beispiel 10: Herstellung von HCV-Vollängen-Konstrukten mit integrierten Fremdgenen zur Verwendung als leberzellspezifische Genfährer für die Gentherapie oder als Expressionsvektor.

[0084] Das Konstrukt (Fig. 14) wird in Zellen eingeschleust und führt dort zur Bildung von HCV-Viruspartikeln, die zur Infektion weiterer Zellen verwendet werden können. Da die Viruspartikel eine RNA mit einem Fremdgen enkapsidiert haben kann dieses in den so infizierten Zellen zur Produktion des von diesem Fremdgen kodierten Proteins benutzt werden. Zellen, die mit dem Konstrukt transfiziert wurden, exprimieren ebenfalls das Fremdgen.

Beispiel 11: Herstellung von monoclonalen HCV-RNA-Konstrukten, bei denen das Resistenzgenprodukt als Fusionsprotein mit dem HCV-Anteil exprimiert wird.

[0085] Für bestimmte Untersuchungen ist es von Vorteil, wenn das HCV-RNA-Konstrukt kein heterologes IRES-Element besitzt. Solche Untersuchungen sind beispielsweise die Bestimmung der Interferonresistenz. Wird eine Zelle, die ein HCV-RNA-Konstrukt besitzt, mit Interferon-alpha oder -beta inkubiert, kommt es zu einer Reduktion der Replikation der HCV-RNA. Zur Aufklärung des Wirkungsmechanismus ist es notwendig, daß das HCV-RNA-Konstrukt keine heterologe IRES besitzt, da ansonsten nicht bestimmt werden kann, ob die Interferon-vermittelte Hemmung durch eine Hemmung der HCV-Replikation oder durch eine Hemmung der heterologen IRES vermittelt wird. Deshalb werden Konstrukte hergestellt, bei denen das Resistenzgen mit einem HCV-Protein fusioniert wird (Fig. 15). Entweder das Fusionsprotein ist als solches aktiv oder das Resistenzgenprodukt wird mittels einer proteolytisch spaltbaren Aminosäuresequenz so mit einem HCV-Protein verbunden ist, daß es von einer zellulären oder viralen (HCV-)Protease von diesem abgespalten wird.

Tabelle 1

Spezifische Infektiositäten (cfu/ μ g RNA) der HCV-RNA-Konstrukte mit adaptiven Mutationen, die bei der 9-13F Mutante gefunden und in das parentale HCV-RNA-Konstrukt 1389/NS3-3'/wt eingeführt wurden		
Aminosäureaustausch ¹	HCV-Protein	cfu/ μ g RNA ²
kein		30 - 60
1283 arg -> gly	NS3	200 - 250
1383 glu -> ala	NS3	30 - 60
1577 lys -> arg	NS3	30 - 60
1609 lys -> glu	NS3	160 - 300
(1283 arg -> gly + 1383 glu -> ala + 1577 lys -> arg + 1609 lys -> glu)	NS3	360 - 420
1936 pro -> ser	NS4B	500 - 1000
2163 glu -> gly	NS5A	1000-5000
2330 lys -> glu	NS5A	30 - 60
2442 ile -> val	NS5B	30 - 60
alle zusammen		5000

¹ Aminosäureaustausch im Polyprotein des HCV-Isolats Con-1 (EMBL-Genbank No. AJ238799); die Aminosäuren sind im Dreibuchstabenkode angegeben.

² Colony forming units (Anzahl der Zellklone) bei einer Selektion von 500 μ g/ml G418.

Tabelle 2

Spezifische Infektiositäten (cfu/ μ g RNA) des parentalen HCV-RNA-Konstrukts I_{329} /NS3-3'/wt und der Varianten 9-13C, 9-13I, 9-13F, 5.1 und 19.		
Transfizierte ante	RNA-Vari-	cfu/ μ g RNA ¹
	Wildtyp	30 - 50
	9-13 C	100 - 1.000
	9-13 I	100 - 1.000
	9-13 F	1.000 - 10.000
	5.1	50.000 - 100.000
	19	50.000 - 100.000

¹ Colony forming units (Anzahl der Zellklone) bei einer Selektion von 500 μ g/ml G418.

Tabelle 3: Nukleotid- und Aminosäureunterschiede zwischen dem parentalen HCV-RNA-Konstrukt I₃₈₉/NS3-3'/wt und den Mutanten 9-13I, 9-13F, 5.1 und 19

HCV Mutante	Nukleotidposition	Nukleotidaustausch	Aminosäureaustausch
9-13 I	3685	C > T	Pro > Leu
	4933	C > T	Thr > Met
	5249	T > C	-
	8486	C > T	-
	8821	G > A	Trp > stop
	8991	C > G	Arg > Gly
	9203	A > G	-
	9313	T > C	Phe > Ser
	9346	T > C	Val > Ala
9-13 F	3866	C > T	-
	4188	A > G	Arg > Gly
	4489	A > C	Glu > Ala
	4562	G > A	-
	4983	T > C	-
	5071	A > G	Lys > Arg
	5166	A > G	Lys > Glu
	6147	C > T	Pro > Ser
	6829	A > G	Glu > Gly
	7329	A > G	Lys > Glu
	7664	A > G	Ile > Val
	8486	C > T	-
	8991	C > G	Arg > Gly
	9346	T > C	Val > Ala
NK5.1	4180	C > T	Thr > Ile
	4679	C > T	-

5		4682	T > C	-
		5610	C > A	Leu > Ile
		6437	A > G	-
10		6666	A > G	Asn > Asp
		6842	C > T	-
		6926	C > T	-
15		6930	T > C	Ser > Pro
		7320	C > T	Pro > Ser
		7389	A > G	Lys > Glu
20	NK19	3946	A > G	Glu > Gly
		4078	C > G	Ala > Gly
		4180	C > T	Thr > Ile
25		4682	T > C	-
		5610	C > A	Leu > Ile
		5958	A > T	Met > Leu
30		6170	T > A	-
		6596	G > A	-
35		6598	C > G	Ala > Gly
		6833	C > T	-
		6842	C > T	-
40		6930	T > C	Ser > Pro
		7141	A > G	Glu > Gly
		7320	C > T	Pro > Ser
45		7389	A > G	Lys > Glu
		7735	G > A	Ser > Asn

50

[0086] Angegeben sind die Unterschiede der Nukleotid- und Aminosäuresequenzen zwischen der Ausgangs-HCV-RNA-Sequenz Con 1 (EMBL-Genbank No. AJ238799) und denen der zellkulturadaptierten HCV-RNAs. Die Zahlen
55 beziehen sich auf die Nukleotid- und Aminosäurepositionen des HCV-Isolats Con1.

SEQUENCE LISTING

5 <110> Baitenschlager, Ralf
 <120> Hepatitis C Virus cell culture system
 <130> ba-1
 <140> 199 15 178.4
 10 <141> 1999-04-03
 <160> 11
 <170> PatentIn Ver. 2.1
 <210> 1
 15 <211> 11076
 <212> DNA
 <213> Hepatitis C Virus
 <400> 1
 20 gcaagccccc gattgggggc gacactcac catagatcac tccctgtgta ggaactactg 60
 tcttcacgca gaaagcgctc agccatggcg ttatgatgag tgtctgcagc cctccaggac 120
 ccccccctcc ggagagagcca tagtggtctg cggaacccgt gactacaccg gaattgccag 180
 gagcagccggg tctttctctg gatcaaccgg ctcaatgcct ggagatttgg gctgcccccc 240
 gcgagactgc tagccgagta gtgttggttc gcgaaggccc ttgtggtact gctgatagg 300
 gtgcttgcca gtgcgccggg aggtctcgta gacctgtcac catgagcagc aatcctaacc 360
 ctcaagaata aacccaacgt aaacccaagc ggogpccat gattgacaa gatgattgc 420
 aocgagttc tccgcgcgct tgggtggaga ggcatctg cytatgactcg gcaacaacaga 480
 caatcgcgct ctctgatgac ccggtgttcc ggctgtcagc gcaggggcgc ccggtctttt 540
 25 ttgtcaagac cgaactgtcc ggtgcctga atgaactgca gacagaggca gcgcgcat 600
 cgttctggcg caacagggcg gttccttgcc cagctgtgct cgaactgttc actgaagcgg 660
 gaagggaactg gctgctattg ggcaagatgc cggggcagga tctctgtcca tctcaacttg 720
 ctcttcgcga gaaagtatcc atcatggctg atgcaatgoc gcgctgcat acgcttgatc 780
 ogctatcccg cccattcgac ccccaagcca aacatgcgat cgaogagaga cgtactcgga 840
 tggaagccgg tcttgctgat cagatgactc tggaagaa gcatcaaggc ctccggcagc 900
 30 ccgaactgtt ccgcaagctc aaggcgcgca tgcccagcgg cgaagatctc gtctgaaccc 960
 atggcgatgc ctgcttgccc aatatcatgg tggaatatgg ccgcttttct gattcaatcg 1020
 actgtggccc gctgggtgtg gcggaacgct atcaggacat agcgttgctt accctgata 1080
 ttgctgaaga gcttgccggc gaatgggctg acogcttccr cgtgctttac ggatcgccc 1140
 ctcccgattc gcaagcgcac gccctctatc gccctctctg caggttcttc tgaatttaa 1200
 cagacccaac cggtttccct ctagcgggat caatcccgcc cctctccctc ccccccctct 1260
 35 aaogttactg gcggaagcgc cttgaaataa ggccggtgtg cgtgttctca tatgttactt 1320
 tcccaatcat tgccgtcttt tgcgaatgtg agggcccgga aaactggccc tctctctctg 1380
 acagcattc ctacgggtct ttcccctctc gccaaaggaa tgcaagttct gttgaatgtc 1440
 gtgaaggaag cagttcctct ggaagcttct tgaagacaaa caactgtctg agcgaacctt 1500
 tgacggcagc ggaacccccc acctggcgac aggtgcctct gcggccaaaa gccactgtta 1560
 taagatacac ctgcaaaagg ggcacaaccc cagtgcacag ttgtgagttg gatagttgtg 1620
 40 gaagagctca aatgtctctc ctcaagcgta caacaaagg gtagtaggga tgcagagaa 1680
 gtacccactt gtagtggctc tctatggggc cctctgtgca catgcttacc atgtgttag 1740
 tcgaagttaa aaacagctta ggccccccca accacgggga cgtgttttct ctttgaaaaa 1800
 caogataata ccatgggcac gaattctaaa cctcaagaaa aaacccaacg taacccaacc 1860
 ogccgccccc aggaactcaa gttcccgggc ggtgttgaag tctgctgggt agtttacctg 1920
 45 ttgcgcgcga ggggcccacg gttgggtgtg ccgcgcacta ggaagacttc cgaoggtctg 1980
 gactccctgt gaagcgacaa acctatcccc aaggctcgcc agcccgaggg tagggcctgg 2040
 gctcagcccg ggtaccctct gccaccttat ggaatgagg gcttgggtg ggaagatgg 2100
 cctctgtcac ccggtgctc tccgctagtt tgggcgcaca cgaacccccc cgttagtctg 2160
 cgcaatttgg gtaagttcat cgtatccctc acgtcgcgct tccgcgatct catgggtcac 2220
 attccctctg tccggccccc cctagggggc gctgcacagg cctcgcgcca tggcctccg 2280
 gttctggagg accgctgtaa catgcaaca gggaatctgc ccggttgcct cttttatctc 2340
 50 tctcttttgg ctttctgtct cttttgacc atcccagctt cccgttatga agtgcgcaac 2400
 gtatccggag tgtacatgt caccagacac tgcctcaacg caagcattgt gtagaggcca 2460
 gcgcacatga tcatgcatac ccccggtgtc gtcocctcgc ttoggagaaa caactccctc 2520
 cgttctgggt tagcgtgac tccagctcgc gcggcgagga accgtatgct cccactacag 2580
 acgtatcgac gccatgtgca ttgtctgttt gggcggtgt cctctgtgct ccgtatgta 2640
 gtggggagat tctggggatc tgttttcttc gtcgcgcagc tgttcaactt ctcgctcgc 2700

	cgccacgaga	cagtcacagg	ctgcaattgc	tcaatatatc	cggccgaacgt	gaacaggtcac	2760
	cytatggctt	gggatattga	gatgaactgg	tcaactacag	caqcccaagt	tatatcgaca	2820
	ttaactccgga	teccacaacg	tgtcgtggat	atggctggcgg	ggcccaacttg	ggaggtcccta	2880
5	cgccgacctg	cctactatct	catggctggg	aactggccta	aggtctctgat	tgtgatgcta	2940
	ctcttttccg	cggttgacgg	gggaacctat	gtgacagggg	ggacgatggc	caaaaacacc	3000
	ctcgggatta	cgctccctct	ttcaccocgg	tcatccocaga	aaatccagct	tgtaaacacc	3060
	aacggcagct	ggcacatcaa	caggactggc	ctgaactgca	atgactccct	caacactggg	3120
	tcccttgctg	cgtcgttcta	ctgtgacaa	tccaactcat	ctgagtggcc	agagcgcatg	3180
	gcacagtgca	ggcccataga	cgggttcctc	caaggttggg	ggctccatca	ttcaactatg	3240
	tcacacagct	cggacacagg	gccttattgt	tggcaattag	caaccocggc	gtgcggtatc	3300
	gtaccccgcg	cgcaggtgtg	tgttccagtg	tactgcttca	ccccaaagcc	tgtcgttgtg	3360
10	gggagcagcc	acccgttcgg	cgctccatcg	tacagttggg	gggagaatga	gacggcagtg	3420
	ctgcttctta	acaacacgcy	gcgcgcgcga	ggcaactggt	ttgctgtcac	atgatgaat	3480
	agcactgggt	tcaccaagac	gtgcgggggc	ccccctgtga	acaatcgggg	gatcgccaat	3540
	aaaaacttga	cctgcgccac	ggactgcttc	cggaaagcac	cggaggccac	ttacaccaa	3600
	tgtggctcgg	ggccttggtt	gacacccaga	tgttggctgc	actaccata	cagcgtcttg	3660
	cactacccct	gcactgtcaa	cttaaccctc	tlcaaggtta	ggatgtacgt	gtggggagtg	3720
15	gagccacgag	tgaagcgcgc	atgcaattgg	actcggaggag	agcgttgttaa	ctcggagagc	3780
	agggacagat	cagagcttag	cccgtctgtc	ctgtctacaa	cggagtggca	ggtattggcc	3840
	tgttctctca	ccacccatac	ggctctgtcc	actggtttga	tccatctcca	tcagaacgtc	3900
	ctggcagctac	aataacctga	cggatatagg	ctggcggttg	tctctcttgc	aatacaatgg	3960
	gagtatgtcc	tgttctctct	ccttctctct	ggcgacgcgc	gcgtctgtgc	ctgcttgtgg	4020
20	atgatgtcgc	tgatgatgca	agctgaggcc	ggcctaaaga	acctggtggt	cccccaacgg	4080
	ctactcgtcg	ccgggggaga	tggcatctct	tccctctgtc	tgtctgtcgt	tgctgctgtg	4140
	tacatcaacgg	gcaggtgtgt	cccctggggcg	gcatactgcc	tccagtcggt	atggcccgct	4200
	ctcctgtctc	tgtctgcgtt	accacccaga	gcatacggca	tgagccggga	gatggcagca	4260
	tcgtgcggag	gcgcgggttt	cgtaggtctg	atacctttga	ccttctgcac	gcactataag	4320
	ctgttctcgc	ctaggctcat	atgggtgcta	caatatatta	tcacacaggc	cggaggcacac	4380
	tgcgaagtgt	ggatccccc	cctcaacgtt	cgggggggcg	gcgatgccct	catcctcttc	4440
25	acgtgcgcga	tcacccacga	gctaattctt	accatcacca	aaatcttctg	cgccatcttc	4500
	ggctccatca	tgggtgcata	ggctgggata	accaaatgct	cgactctcgt	gcgcgcacg	4560
	gggtctcatc	gtgcactgat	ctctgtgcgg	aaggttgctg	gggtctctgt	tgtccaaatg	4620
	gctctcatga	agttggcgcg	actcgaaggt	acgtacgttt	atgacatctc	caccccaatg	4680
	cgggacatgg	cccacgcggg	cctacgagac	cttgcggtgg	cagttgagcc	cgtcgtcttc	4740
	cttgatgatgg	agacaaaggt	tatacaactg	ggggcagaca	ccgcgcggtg	tggggacatc	4800
	atcttggggc	tgcctctctc	cgccgcgcag	gggagggaga	tacatctggg	accggcgagc	4860
30	agccttgaag	ggcagggggtg	gcgactcttc	cgccctatta	cgcctacttc	ccaacagacg	4920
	cggaggccac	ctggcctcat	ctacaagtcg	gggacaggaa	ccagctcgag	acagctcgag	4980
	ggggaggtcc	aagtggtctc	caacgcacaa	caatctttcc	tgggcacgtg	cgtccaatgg	5040
	gttgtgtgga	ctgtctatca	tgtgtccgcg	tcaaaagacc	ttgcggccgc	aaaggcccca	5100
	atcaccacaa	tgtacaccaa	tgtggacag	gacctctcgc	gctgcgaacg	gccccccggg	5160
	gcgcgttctc	tgtacacatg	caactcgcgc	agctcggacc	tttaacttgt	cacagagcat	5220
	gcgcagtgtca	ttccggtgcg	ccggcggggc	gacagcaggg	ggagcctact	ctccccacgg	5280
35	cccgtctctc	acttgaaggg	ctcttcgggc	ggtccactgc	tctgcccttc	ggggcacgct	5340
	gtggccatct	ttcgggtgcg	cgtgtcaccc	caggggggtg	cgaaggcggt	agactttgta	5400
	ccgtctcagc	ctatggaaac	cactatcggt	tcccggctct	tcaggacaaa	ctcgtccctc	5460
	ccggcgctac	cgcagcaact	ccaggtggcc	catctacacg	ccctactcgt	tatcgccaaa	5520
	agcactaaag	tgcgcgtcgc	gtatgcagcc	caagggtata	aggtgtctgt	cttgaacccc	5580
	tcctgtccgc	ccacccatag	tttcggggcg	tatatgtcta	agggcacatg	tatcgacctc	5640
40	aaactacgaa	ccggggtaag	gacctacacc	acgggtgccc	ccatcacgta	ctccacttat	5700
	ggcaagtctc	ttgcgcagcg	tgtgtgctct	gggggcgcct	atgacatcat	aatatgtgat	5760
	tggtgccact	caactgactc	gacctatc	ctgggcacgc	gcacagctct	ggaccaagcg	5820
	gagacggctg	gagcgcagat	cgtctgtctc	ggcacgcgta	cgcctccggg	atcgtctcac	5880
	gtgcacatca	caacatgaga	ggagtgaggt	tgttcacaga	ctggagaaat	ccccctttat	5940
	ggcaaaagca	tecccatcga	gacctacaa	ggggggagcg	acctcaattt	cttgcattcc	6000
	aagaagaat	gtgatgagct	cgcgcggaag	ctgttcgggc	tcggacatcaa	tgtcttagca	6060
45	tattacccgg	gcctttgatg	atccgtcata	ccaactagcg	gagacgtcat	tgtcgtagca	6120
	acggacgcct	taa'gacgcy	ctttacccgc	gatttcagct	caytgatoga	ctgcaatata	6180
	tgtgtccacc	agacagtcga	cttcagcctg	gaacccgact	tcacatttga	gacgcagcac	6240
	gtgcacaaag	acccgggtgc	acgtcgtcag	cggcgaggca	ggaactggtag	gggcgggatg	6300
	ggcatctaca	ggtttgtgac	tcaggagaaa	cggccctcgg	gcattgttca	ttccctggtt	6360
	ctgtgcagat	gcataagaga	ggagtgaggt	tgttgcagga	ctacgcgcgc	cayagactca	6420
	ctgaggttgc	ggccttaact	aaacacacaa	gggttgcgcg	tctgcacaga	ccactctgag	6480
50	ttctgggaga	gcctttttaa	aggctccacc	cacatagacg	ccattttctt	gtccacgact	6540
	aacagcgacg	gagacaactt	ccccctactg	gtagcatacc	aggtcacggt	gtgcgccagg	6600
	gcctcagctc	cactctccat	gtgggaccaa	atgtggaagt	gtctctacgc	gctaaagcct	6660

agcgtgcacg ggcacaagcc cctgtgtgat agcgtggag cggttcaaaa cagagttact 6720
 accacacacc ccaataccaa atacatcatg ccatcatgt cggctgacct cagagtgctc 6780
 acgagcacct ggggtgttgt aggcgcagtc ctacagcttc tggccgcgta ttgectgaca 6840
 acaggcagcy tggctcattg ggcgcagatc atcttgtccg gaaagccgcg atcatctccc 6900
 gacagggaaq tccctttacc ggaagtccat gagatggaaq agtgcgcctc acacctccct 6960
 tacat ogaac agggaatgca gctgcgcgaa caattcaaac agaagccmat cgggttgtgt 7020
 caaacagcca ccaagcaagc ggaagctgct gctcccgctg fgggaatccaa gtgcgggaac 7080
 ctcpagacct tctcggcgca gcatatgttg aatttcaaa agtggatcaca atatttagca 7140
 ggcgtgtgca ctctcctgt caacccgcg atagatcac tgaatgact caacgctct 7200
 atcacagccc cgtctaccac ccaacatacc ctctgttta acatctggg ggaagtgggt 7260
 gccgcaccaac ttgtctctcc cagcgtgctc ttctgtttcg taggcgcgcg catctgtgga 7320
 gcgcgtctgt gcagcatagg ccttgggaag gtgctgtgtg atattttggc aggttatgga 7380
 tccagctgtg cagcgcgctc cgtggccttt aaggtcaatga gcgcgcagat gccctccacc 7440
 gaggacctgg ttaacctact cctctccctg ggcgcctagt cgtcggggtc 7500
 ggtgtgcagc cgaatactgc tggccagctg ggcocaggag agggggctgt cgaagtgtgt 7560
 aacgtgtgca taggttctgc ttcgcggggt aacacgtct cccacagcta ctatgtgct 7620
 gagagyaagc ctgcgcgacg tgcctccag atcctctcta gtcttaccat cactcagctg 7680
 ctgaagagcc ttcacacagt gatcaacagc gactgctcca cgcctctctc cgtctgtgga 7740
 ctaagagatg ttgtggattg gatagtcacg gtgttgactg atttcaagac ctggctccag 7800
 tccatagctcc tgcgcgcatl gccggagatc cctttctct catgtcaacg tgggtacaag 7860
 gga gctgtcgc ggggcgcagc catcatgaca acaactgcc catgtggagc acagatccgc 7920
 ggaactgtga aaaaaggctc catgagatc ttggggcccta ggaactgtag taacagctgt 7980
 tgtgtgaact tctgtgctta cgcgcacacc cagggccctt gcaagccttc cccgggccta 8040
 aattattcta aggcgctgtg gcgggtgctg cgtcgagagt cgtcgagagt cagcagctg 8100
 gggagatttcc actacgtgac gggcgtgacc actgacacgc taaagtcccc gtgtcaggtt 8160
 ccgcgccccg aattcttacc agaagtggat ggggtgcggt tgcacagcta cgtctcaagc 8220
 tgcacaaccc tctactggga ggaagtgaca tctctgtct ggcctcaatca atacctggtt 8280
 gggctcacagc tcccatgga gcccgaaagc gaagttagag tgcctcactc catgtccacc 8340
 gacacctccc acattacgac ggagacgctc aagcgtaggc ttggccagggg atctcccccc 8400
 tcccttgcca gctcatagc tagccagctg tctgcgcctt ccttgaaagc aacatgacct 8460
 acccgctcag actccccga cgcgtgauct atcgaggcca acctctgtgt ggcgggtgag 8520
 atggcgggga acatcacccc cgtggagtca gaataaagp tagtaatttt gactcttttc 8580
 gagccctccc aagcggagga ggaatgagag gaagtatacc ttcocggogga gatcctgcg 8640
 aggttcagga aattccctcg agcgaatgcc atatgggacc gccocgatta caacctccca 8700
 ctgttagagt cctgggaagga cccggaactac gtccctccag tggtaacagc gtgtccattg 8760
 ccgctgcgca aggcocctcc gataccaact ccaaggagga agaggaoggt tgtcctgtca 8820
 gaactacccg tgcctctctc ottggcgagc ctgcgcacaa agactcttgg cagctccgaa 8880
 tctgtgcgac tgacagcgag caggcgagc gctcctctg acagccttc cgaagcagcg 8940
 gacggcgatc ccgacgttga gtcgtactcc tccatgcgcc ccttgagggg ggaagcggg 9000
 gatccccctc tcagcgacgg gtctttgtct acgttaagc aggaagctag tgagcagctc 9060
 gtctgtctct cgaatctcta cacatggaca ggcgcctga tcaagcctat cgtcggga 9120
 gaaccaagc tgccatacaa tgcactgagc aactctttgc tccgtcacca caacttggtc 9180
 tatgtacaaa catctgcgac cgcaagcctg cggcagaaga aggtcaacct tgacagact 9240
 caggtcctgt acgacacta cccgggcagct ctcaaggaga tgaagcgcaa ggcgtccaca 9300
 gttaaagcta aactctatc cctggaggaa cctgttaagc tgaagccccc acatcgccc 9360
 agatcaaat ttgctatgg gcgcaagac tgcgggaa taltccgaaa ggcgtttaa 9420
 cacatcgctc ccgigtgga ggaactgtct gaagacactg agacacaa tgacacacc 9480
 atcatggcaa aaatagagt ttctcgtc caaccagga agggggggcc caagcagct 9540
 cgccttatcg tattcccaag ttggggggtt cgtgtgtgcg agaaaatggc cctttacgat 9600
 gtgtgtccca cctccctca ggcgctgat ggccttcat acggaatcca atactctct 9660
 ggaagcgggg toagtttct ggtgatctc tggaaagca agaatgccc tatgggcttc 9720
 gttatgaca cccgctgttt tgactcaacg gtaactgaga atgacatcg ttttgagag 9780
 tcaatcacc aatgtgtga ctggccccc gaagagaa aggcataagc gtgcctcaa 9840
 gagcgctt caatcgagg cccctgact aattcttaa ggcgaactg ggcgaactg 9900
 cgglgcgcgc cpagcgggtt actgacgacc agctcggtta atacctcac atgttactg 9960
 aagcgcctgt cggcctgtct agctcggaag ctccaggact gcaacatgct cgtatgogga 10020
 gaogaactgt tcttatctg tgaagacgcy gggaccccaa aggaagagcg gagctacg 10080
 ccttccagc aggcctatgac tagatactct gcccccctg gggacccgoc caaacacaaa 10140
 tacagacttg agttgatac atcatgctc tcaaatgtg cagtcogca cgtatcatc 10200
 ggcacaaagg tctactatct caccgctgac cccacacacc ccttgcgagc ggtcgtgtg 10260
 gagacagcta gacactctc agtcaatcc ttgctaggca acatacat gtatgcgcc 10320
 acctgtgg acctcgagg cctgagact catctctat ccatctct agtcagaa 10380
 caacttga aaagcctaga ttgtcagatc taaggggct gttactcaat tgaacacct 10440
 gactcactc agatcattca acgactccat ggccttagc catlttact ccaatgtta 10500
 tcttcaggtg agatcaatag ggtgcttca tgcctcaac aacttgggt accgcctgt 10560
 cgaattcga gacatgggc cagaaggttc cgcgtaggc cggggggag 10620

EP 1 043 399 A2

5 gctgccactt gtggcaagta cctcttcaac tgggcagtaa ggaccaagct caaactcact 10680
 ccaatcccgg ctgcgtccca gttggattta tccagctggg tcgttgctgg ttacagcggg 10740
 ggagacatat atcacagcct gtctcgtgcc cgaccccgct ggttcatgtg gtgcctactc 10800
 ctactttctg taggggtagg catctatcta ctccccaacc gatgaacggg gagctaaaca 10860
 ctccaggcca ataggccatc ctgttttttt cccttttttt tttctttttt tttttttttt 10920
 tttttttttt tttttttttc tccttttttt ttctcttttt tttccttttc tttccttttg 10980
 tggctccatc ttagccctag tcacggctag ctgtgaaagg tccgtgagcc gcttgactgc 11040
 agagagtgct gatactggcc tctctgcaga tcaagt 11076
 10
 15
 20
 25
 30
 35
 40
 45
 50
 55

<210> 2
 <211> 8637
 <212> DNA
 <213> Hepatitis C Virus

<400> 2
 gccagccccc gattgggggc gacactccac catagatcac tccccgtga ggaactactg 60
 tcttccagca gaagcgctct agcactggcg ttagtatgag tgcgtgcag cctccagpac 120
 cccctctccc gggagagca taaaggctct cggaaacgg gaaatgcag 180
 gacgacggg tctctctctg gatcaacccc ctcaatgcct ggagattctg gctgcccccc 240
 gcgagactgc tagcccgata gtgttggttc gcgaaggccc ttgtggtact gctgtatagg 300
 gtgcttgcaa gtgccccggg aggtctcgtta gacctggcac catgacagac aatcctaaac 360
 caagaaagaa aaccaaaggc ggcgcactga ttgacaagaa tgaattgcac gcaggtctct 420
 cggccgcttg ggtggagagg ctattcggct atgactgggc acaacagaca atcgctcgtc 480
 ctgatgccgc cgtgttccgg ctgtcagcgc agggcgcccc ggtctctttt gtcaagaccg 540
 acctgtccgg tgcctgaat gaactgcagg acgagtcagg ggggctatcg tggctgcaca 600
 cgaaggcggt tctctgcaga gctgtgctcg acgtgtccac tgaagcggga agggactcgc 660
 tgcattatgg cgaagtgcgg ggcgaagctc tcccttcact tcaacttctc cctgcgagaa 720
 aagtatccat catggctgat gcaatgcggc ggtgcatac gcttgatcgc gctacctgcg 780
 cattgcacca ccaagcgaag catcgcatcg agcgagcacg tactcgtatg gaagccgctc 840
 ttgtcgatca ggaatgatctg gaagaaagac atcaggggct cgcgcagacc gaactgttcg 900
 ccaagctcaa ggcgcgcgat ccgcagcgcg aggaatcgtt cgtgaacctt ggcgatcgct 960
 gctgtccgaa tatcatgggt gaaaatggcc gctttctcgt atcatcgac tgtgcccgcg 1020
 tgggtgtggc ggaccgctat caggacatag cgttggctac cctgtatatt gctgaagagc 1080
 ttggcgcgga atgggtgag cgtcttcacg tgccttaccg tatcgccgtc ccgactgcgc 1140
 agcgactcgc ctctcatcgc ctctcttcagc agttcttcgt agtttaaaac gacccaacag 1200
 gtttccctct agcgggatac attcgccccc tctccctccc cccccctaa cgttactcgc 1260
 cgaagccgct tggaaatagg ccggtgtgccc tttgtctata tgtttatttc caccattatt 1320
 ccgctctttt gcaatgtgag ggccccgaaa cctggccctg tctctctgac gacattctct 1380
 aggggtcttt cccctctcgc caaagggaat caaggtctgt tgaatgtcgt gaaggaaaga 1440
 gttcctctgg aagctctctg aagacaacaa acgtctgtag cgaacctttg caggcagcgg 1500
 aacccccacc ctggcgacag gtgcctctcg gcgcaaaagc cacgtgtata agatacacct 1560
 gcaaggcgcg caaaccccca gtgccacgtt gtgagttgga tagtgttga aagatgcaaa 1620
 tggctctctt caagcgtatt caacaaggcg ctgaaggatg ccacagaagt acccaattgt 1680
 atgggactcg atctggggcc tctgtgcaca tgccttacct gtgtttagtc gaggttaaaa 1740
 aacgtctagg cccccgaac caccgggagc tggtttccct ttgaaaaaca cgataatacc 1800
 atggaccggg agatggcagc atcgtgcgga ggccgggttt tctgtagtct gatactcttg 1860
 acctgtgcac cgcactataa gctgtctccc gctaggctca tatgtgtctt acaattttt 1920
 ataccagggg ccgagcgaca ctgtcaagtg tggatccccc cctcaacgtt tcggggggggc 1980
 ccgatgcgcg tcatctctct cacgtgcgcg atccaccag agctaacttt taccatcac 2040
 aaatctctgc tgcgcatact cgtgtcactc atgggtctcc aggcgtgatt aacccaagtg 2100
 ccgtactctg ttgcgcgaca cgggtctcat cgtgcaagca tgcgtgtgag gaaggttctg 2160
 gggggtcaatt atgtccaaat gctctctatg aagttggcgc catgacagc tacgtacctg 2220
 tatgaccatt tcaccccact cggcgactgg gccaccggcg gctacacaga ccttgccggt 2280
 gcagtgtgac cgtctgtctt ctctgatgat gagaccaagg ttatccactt gggggcagac 2340
 accgcggcgt gtggggacat catcttgggc ctgcccgctc ccgcgcagag ggggagggag 2400
 atacactctg gaccggcaga cagccttgaa gggcaggggt ggcgactcct cgcgcctatt 2460
 acgcccactt cccaacagag cggaggccta ctgtgctgca tctacactag cctcacagcg 2520
 cgggacagga accaggctga gggggaagtc caagtggctt ccacgcgaac acaattcttc 2580
 ctggcgacct gctgcaatg cgtgtgtgtg actgtctatc ataggtcggg cccaagagac 2640
 ctctgcggcc caaaggccct aatcaacca atgacacca atgtgagaca ggactctgc 2700
 ggtctgcagg gcgccccggg ggcgcttccc ttgaacacct gcactcggcg cagctcgagc 2760
 ctttactctg tcaagaggca tgcgcagtgc attccggtgc gcgcgcgggg cgacagcagg 2820
 gggagcctac tctccccag cgcgctctcc tacttgaagg gctcttcggg cgttccactg 2880
 cctcgcacct cggggcagcg tgtggggcac ttccgggtcg ccgtgtgcac ccgaagggtt 2940
 gogaaggcgg tggactttgt acccgtogag ttatatgaaa cccatctagc gtccccgctc 3000
 ttacggacca actcgtcccc tccggccgta ccgcagacat tccagtggtc ccactctaac 3060
 ggcctcatct gtaagcgaaa ggcgactaa gtccggctcg cgtatgcag ccgaaggtat 3120
 aaggtgcttg tctgaaaccc gtccgtcgcc gccaccctag gtttcggggc gtaatatgtc 3180
 aagcacaatg gtatgcaccc taactacaga accgggttaa ggaccatac caccgggtcc 3240
 cccatcaagt actcaacata tggcaagttt cttgcgcagc gtggtgtctc tggggcgccc 3300
 tatgacatac taatatgtga tgagtgcac tcaactgact cgaccactat cctgggcatc 3360
 ggcacagtcc tggaccaagc ggagacggct ggagcgagac tctgtcgtcg gccaccgct 3420
 accgctccgg gatcgttac cgtgccatat ccaacatcg aggaagtggc tctgtccagc 3480
 actggagaaa tcccccttta ttgcaaaagcc atccccatcg agaccataa gggggggagg 3540

	caactcattt	tctgcattc	caagaagaa	tgtgatgagc	tgcgcgaa	gctgtccggc	3600
	ctcgactca	atgtgtgagc	atatccgg	ggccttgatg	tatccgtcat	taaccatagc	3660
	ggagacgtca	ttgtctgtac	aacggacgtc	ctaatacgagc	gctttacacg	cgattttcagc	3720
5	tcagttgatc	actgcatac	atgtgtcaac	cgacacagtc	acttcagcct	ggaccgcagc	3780
	ttccaccattg	agacgcagac	cgtgccacaa	cgccgggtgt	caacgttcga	cgccgcagagc	3840
	aggaactgta	ggggcagat	gggcatttac	aggttttgga	ctccaggaga	acggccctcg	3900
	ggcatgttcc	attcctcggt	tctgtgcgag	tgcattgacg	cgggctgtgc	ttggtacagc	3960
	ctcagccgcc	cgagacactc	agttaggttg	ggggctaac	taaacacac	aggttgccc	4020
	gtctccagc	acactctgg	gtctaggag	agcgtctca	caggctcac	ccacataaac	4080
	gccattttc	tgtcccgagc	taagcagcca	ggagacacat	tcctccactc	ggtagcatac	4140
	cgaggtacgg	tgtgcgcagc	ggctcagcgt	ccactccat	cgtgggacca	aatgtggaag	4200
10	ttgtctcatac	ggctaaagcc	tacgtcgac	gggccaacgc	ccctgtcgtg	taggttgagg	4260
	gcggttcaaa	acgaaggttac	taccacacac	cccataacca	aatacatcat	ggcatgcatg	4320
	tcgggtgacc	tgaaggtcgt	cacgagcaac	tgggtgctgg	tagggcgagt	cttagcagct	4380
	ctggccgcgt	attgctgac	aacaggcagc	gtggtcattg	tgggcaggat	catcttgccc	4440
	ggaaagccgg	ccatctatcc	cgacaggga	gtccctaac	ggaggttcga	tgaatggaa	4500
	gagtcgcct	ccacactccc	ttacatcgaa	cagggaatgc	agctcgcca	ccaatccaa	4560
15	cgaaagcna	tccggttgct	gcacacagc	accaagcaag	cggaggctgc	tgctccctgt	4620
	tggaatacca	agtgcgcagc	ctccgaagcc	ttctggcgca	agcatatgtg	gaatttcac	4680
	agcgggatac	aatatttagc	aggtctgtcc	actctgcctg	gcaacccccc	gatagcatca	4740
	ctgatggcat	tcacagcctc	tatccacagc	ccgctcacca	cccaacatac	ctcctgtttt	4800
	aaactcctgg	ggggatgggt	ggccgcocaa	cttgctccct	ccagcgtcgc	ttctgcttcc	4860
20	gtagccgcgc	tgactcgtgg	agcgctgtgt	ggcagcatag	gctctggaaa	ggtgcttgcc	4920
	gaattttcga	ccgttatgg	agcgaagtg	cgacgcgcgc	tctgtgcctt	taagatcgat	4980
	ggcggcgaga	tgcctccccc	cgaggtctgt	gttaacctac	tccctgtctc	ctctctccct	5040
	ggcgccctag	ctctcggggt	cgtgtgcga	gcgatactgc	gtccgcacgt	gggcccaagg	5100
	gagggggctg	tgcagtgat	gaacggcgtg	atagcgtctg	cttcgcgggg	taaccacgtc	5160
	tcocccacgc	actatgtgcc	tgaagcgac	gctgcagcac	gtgtcaactca	gatctctctc	5220
	agctctacca	tcaactcagct	gctgaagagg	cttcaccagt	ggatcaacga	ggactgctcc	5280
25	acgcacatgt	ccggctctgt	gctaaagat	gtttgggatt	ggatatgccc	ggtgttgact	5340
	gatttccaga	ccgtgctca	gcacacgtc	ctgcggcgat	tggcccgact	ccctctcttc	5400
	tcattgtcac	gtggtacaa	gggggtctgt	cggggcgcgc	gcatcatgca	aaccacatgc	5460
	ccatgtggag	ccacgatac	cggacatgtg	aaacacggt	ccatggagat	cgtggggcct	5520
	aggaactgta	gtaacacgtg	gcattggaaca	ttcccccatt	acgggtacac	cacgggcccc	5580
	tgccagccct	cccgcgcgcc	aaattattct	agggcgctgt	gcgggtgccc	tgctgaggag	5640
	taactgtgag	ttacgcgggt	gggggatttc	cactacgtga	cgggcatacg	cactgacac	5700
30	gttaaagtcc	cgtgtcaggt	tccgcgccccc	gaattcttca	cagaagtga	tggggtgcgg	5760
	ttgcacaggt	agctctcagc	gtgcaaaccc	ctctacaggy	aggaagttac	attcttggtc	5820
	ggggtcaatc	aatactcgtt	tgggtcacag	ctccatggg	agccgaaccc	ggaactgaca	5880
	gtgctcaatt	ccatgtctac	cgcctccctc	ccattarccg	cggagacgcg	taagcttagg	5940
	ctggccaggg	gatctccccc	ctccttgccc	agctcatcag	ctagccagct	gtctgcgctc	6000
	tccttggaag	caacatgcac	taccgcctat	gaactcccgc	acgctgacct	catcgaggcc	6060
	aacctcctgt	ggcgccagga	gatgggcggg	aacatcaccc	gcgtggagtc	agaaaataag	6120
35	gtgataattt	tggactcttt	cgagccgcctc	caacggaggg	aggaatgag	ggaagtatcc	6180
	gttccgcggg	agatctctgc	gaggtccagg	aaattccctc	gagcgatgcg	catatgggca	6240
	cgccgggatt	acacactccc	actgttagag	tccgtgaagg	accgggacta	ctgccttcca	6300
	gtggtcacag	ggtgtccatt	gcgcctgcgc	aaggccctcc	cgataccacc	tccacggggc	6360
	aagaggacgg	ttgtctctgc	agaactcac	gtgtctctg	ctttggcgga	gtctgcacca	6420
	aagaccttcg	cgacgtccga	atcgttcggc	gtcgacagcg	gcaacggcac	ggcctctcct	6480
40	gaccagccct	ccgacgcagg	cgacgcggga	tcgcagcttg	agtgcacttc	ctccatgccc	6540
	ccctctgaag	gggacgcggg	ggatcccgat	ctcagcagag	ggtcttggtc	taccgttaagc	6600
	gaggaggcta	gtgaggacgt	cgtctgtcgc	togatgtcct	acacatggac	agggccctcg	6660
	atccagccat	gcgctgcgga	ggaaaccaa	ctgcaccata	atgcactgag	caactctttg	6720
	ctccgcaccc	acactctctc	ctatgcagaca	acactgcgca	ggcgagacct	ggcgagagc	6780
	aaggtcacct	ttgacagact	gcaggtccgt	gacgacccat	accgggacgt	gctcaaggag	6840
	atgaagcgga	agggotccac	agttaaagct	aaactcttat	ccgtggagga	agcctgttaag	6900
45	ctgacgcccc	cacattcggc	cagatctaaa	tttgctatg	gggcaaaagg	cgtccgggac	6960
	ctatccagca	agggcgttaa	ccacatccgc	tcogtgtgga	aggaacttgt	ggaagacact	7020
	gagacaccaa	ttgacaacca	catcatggca	aaaaatgagg	ttttctgcgt	ccaacagag	7080
	aaggggggcc	ccaagccagc	tcgccttatt	gtattccag	atttgggggt	tcgtgtgtgc	7140
	gajaaaatgg	cccttttaca	tgtggtctcc	acctccctcc	agccgctgat	gggctcttca	7200
	tacggatccc	aatctctctc	tcgacagggg	gtcgacttcc	tggttgaaag	ctgacaagcc	7260
	aagaattgoc	ctatgggctt	cgtatataac	accgctgctg	ttgactcaac	ggctactgag	7320
50	aatgacatcc	gtgttgaggga	gtcaatctac	caatgtttgt	actttggccc	cgagccagga	7380
	caggccacaa	ggctcgtcac	agagccgctt	tacatccggg	ggccctcgac	taattctaaa	7440
	gggcagaact	cggtcatctg	ccggtgcgcg	ggagcgggtg	tactgacgac	cagctgcggt	7500

aataccctca catgttactt gaaggccgct gggcctgtc gagctgcgaa gctccaggac 7560
 tgcacgatgc tcgtatgcgg agacgacctt gtcgttatct gtgaaagcgc ggggacccaa 7620
 5 gaggacgagg cgagcctacg ggccttcacg gaggctatga ctagatactc tgccccccct 7680
 ggggacccgc ccaaaccaga atacgacttg gagttgataa catcatgctc ctccaatgtg 7740
 tcagtcgcgc acgatgcac tggcaaaagg gtgtactatc tcacccgatga cccaccacc 7800
 ccccttcgcg gggctgcgtg ggagacagct agacacatc cagtcaatc ctggctaggc 7860
 aacatcatca tgtatgcgcc caccttgtgg gcaaggatga tcctgatgac tcatttcttc 7920
 10 tccatctctc tagctcagga acaacttgaa aaagccctag attgtcagat ctacggggcc 7980
 tgttactcca ttgagccact tgacctacct cagatcattc aacgactcca tggccttagc 8040
 gcattttcac tccatagtta ctctccaggt gagatcaata ggggtggcttc atgcctcagg 8100
 aaacttgggg tacgcacctt gcgagtcctg agacatcggg ccagaagtgt ccgcgctagg 8160
 ctactgtccc agggggggag ggctgccact tgtggcaagt acctctcaa ctgggcagta 8220
 aggaccaagc tcaaacacac tccaatcccg gctgcgtccc agttggattt atccagctgg 8280
 15 ttctgtctgt gttacagcgg gggagacata tatcacagcc tgtctcgtgc ccgaccgcgc 8340
 ttgttcatgt ggtgcctact cctactttct gtaggggtag gcacttatct atccccaac 8400
 cgatgaacgg ggagctaaac actccagycg aataggccat cctgtttttt tccctttttt 8460
 ttttcttttt tttttttttt tttttttttt zttttttttt ctcccttttt ttctcttttt 8520
 ttttcttttt ctttctttt gtggctccat cttagcccta gtcacggcta gctgtgaaag 8580
 gtcggtgagc cgcttgactg cagagagtgc tgatactggc ctctctgcag atcaagt 8637
 20
 25
 30
 35
 40
 45
 50
 55

<210> 3
 <211> 8001
 <212> DNA
 <213> Hepatitis C Virus

5

<400> 3
 gccagccccc gattgggggc gacactccac catagatcac tccccttgta ggaactactg 60
 tcttcaogca gaaagcgctc agccatggcg tttagtatgaq ttgtogtcag cctccaggac 120
 ccccccctccc gggagagcga tagtgctctg cgganccoygt gagtacaccc gaattgccag 180
 gacgacccgg tctctttctg gatcaacccg ctcaatgctt ggaagtattg gcgtgccccc 240
 gcgagactcg tagccagta gtgttggttc gcgaaaggcc ttgtgtactt gccctatagg 300
 gtgcttgcga gtgccccggc aggtctcgtta gaccctgcac catgagccag aatcctaacc 360
 ctcaaaagaa aaccaaagct aacccaacgg gcgcgcgcac gattgaacaa gatggattgc 420
 acgcaggttc tccggccgct tgggtggaga ggctattccg ctatgactgg gcacaacaga 480
 caatccgctg ctctgatgcc gcgctgttcc ggctgtcaag gcaggggcgc ccggttcttt 540
 ttgtcaagac cgaactgtcc ggtgcctga atgaactgca ggaagaggca gcgcgcctat 600
 cgttgctggc cagcaagggc gttcccttgc agctgtgtct cgaactgttc actgaagcgg 660
 gaagggaact gctgtctatg gggaaagtgc gcggggcaga tctcctgcga tctcaacttg 720
 ctctccgcga gaaagtatcc atcaatgctg atgcaatgag gcggctgcat acgcttgatc 780
 ccgctactctg cccattogac caccaaagca aacatcgcat cgaagcagca cgtactcga 840
 ttgaaagccgg tcttctgat caggatgtat tggacgaaga gcatcagggg ctccgcgccg 900
 gcgaactgtt ccgcaagctc aaggccgcga tgcccagcgg aggagatctc gtctgacccc 960
 atggcagatc ctgcttccgg aatatcaatg tggaaaatgg ccgctttctt ggaatcatcg 1020
 actctggcgc gctgggtgtg gcggaacgct atcaggacat agcgttggct accoytgata 1080
 ttgctgaaga gcttggcggc gaattggctc aacgcttctc cgtgctttac ggtatgcgcg 1140
 ctcccgatcc gaaagcctc gctctctatc gctctcttga cgaagtcttc tgaagtataa 1200
 cagacaacaa cggtttccct ctacggagat caattccggc cctctccctc ccccccctct 1260
 aacgttactg gccgaagccg cttggaataa ggcccggtgt cgtttgtcta tatgttattt 1320
 tccaccatat tgcctctctt tggcaatgtg agggcccoga aacctggccc tgtctttct 1380
 acgagcattc ctacgggtct ttcccctctc gccaaaggaa tgaagactgt gttgaatgtc 1440
 gtgaaggaag cagttctctt ggaagctctt tgaagaacaa caacgctcgt agcgaacctt 1500
 tgacggcagc ggaaccccccc acctggcgac aggtgcctct gcggcccaaa gccacgtgta 1560
 taagatacac ctgcaaggcg ggcaacaacc cagtgcacgg ttgtgagttg tagattgtgt 1620
 gaaagatca aatgtctctc ctcaagctga tctcaacaaq ggctgaagga tgcacgaag 1680
 gtaccccatt gtatgggatc tgatctgggg cctcggtgca catgctttac atgtgtttat 1740
 tcgaggttaa aaaaactcta ggccccoga accacgggga cgtggtttc ctttgaataa 1800
 cagataata ccatggcgcc tattaccgac tactcccaac agacggcgag cctacttgcc 1860
 tgcatcatca ctacgctcac aggcggggac aggaaccagg tcgaggggga gttccaagt 1920
 gtctccaccg caacaacatc ttccctggcg acctgcgtca atggcgtgtg ttggaactgc 1980
 tatcatgggt gcggtcctaa gaccttgcg ggcccgaagg gcccaatcac ccaaatgtac 2040
 accaatgtgg accaggaact cgtcgggtgc caagcgcgcc ccggggcgcg tctcttgaca 2100
 coactgcaact gggcgagctc ggaactttac ttggtcaaga ggcatggcga tgtcatctcg 2160
 gtgcgcgcgc gggcgacagc caggggagag ctactctccc ccaggccgct tctcactgt 2220
 aagggctctt cgggcggtcc actcgtctgc cccctggggc acgctgtggg catcttctcg 2280
 ggtgcctgtg gccacggagg ggttgogaag gcggtggact ttgtaccctg cgaacttatg 2340
 gaaaccacta tgcgctcccc ggtcttcaag gacaactcgt cccctccgcg cgtaccgcag 2400
 acattccagg tgcgccatct acacgcccct actgttagcg ccaagagcac taagtgccc 2460
 gctgcgtatg cagccogaag gtataaggtg ctgtctctga acccgtccgt ccgcgccacc 2520
 ctagtgttgc gggcgtatat gtctaaagca actgatctcg accctaact cagaacccgg 2580
 gtaaggaoca taaccaaggg tgcccacatc acgtactcca cctatggaa gttctctgc 2640
 gaagcggttt gctctggggg cgcctatgac atcaaatat gtgatgaatg ccaactcaat 2700
 ctacgcacca ctatctctgg catcggaaca ctgctgaacc acgtcgagac ggctgagacg 2760
 cgaactgtcg tgcctgcgac cgtcacgcct ccgggtatcgg tcaacctgcc acatcaaac 2820
 atcagagagg ttgctctgtc cagcaactga gaaatccctt ttatatgcaa agccatcccc 2880
 atcagagcca tcaagggggg gaggcaactc attttctgc attccaagaa gaaatgtgat 2940
 gactgcgcgc cgaagctgtc cgcctctgga ctcaatgctg tagcatata cggggcctt 3000
 gatgtatccg tcatacaac tagcagagac gtcaattgtcg tagcaacoga cgtctaatg 3060
 accggtctta ccgggatctt cactctatgt atcgactgca atacatgtgt cacccaagca 3120
 gtgcactca gctctgaccc gactctacac attgagaca cgaacgtgc caaagagcg 3180
 gtgtcacgct cgcacagcgg agccaggaact ggttagggga ggaatgggat ttaacagttt 3240
 gactctcaga gaaagcggcc ctccggcatg ttccgattct cgtttctgt cgaatgctat 3300
 gacgcgggct gtgcttgata cgaagctcaag cccgcgagaa cctcagttag gttcgggct 3360
 tacttaaaaa cacaaggttt gccctctctg caggacaact tggagttctg ggaagcgtc 3420
 tttaacggcc taccacacat agacgcccat tcttgttccc agactaagca ggcagagac 3480
 aaactccccc acctggtagc ataccaggct accggtgtgc ccagggtctca cgtcccaact 3540

55

	ccatctgtggg	accaaagtgt	gaagttgtctc	atacggctaa	agctacgtct	gcaaggccca	3600
	acgcctctgt	tgttaagctt	ggagtcgctt	caaaacagag	ttactatcac	aaacoccca	3660
	accaaataca	tcatgtgatg	catgtcgctt	gaactggagg	tgcgtacagag	caactcggtg	3720
5	ctgtgtaggcg	gaagtcctagc	agctctggcc	gcgtattggc	tgacaacagg	caagcttggtc	3780
	attgttgggca	ggatcatctt	gtccggaaag	ccggcatca	ttcccagacg	gggaagtcct	3840
	taccggggagt	tcatagagat	gaaagagtcg	gcctcacacc	ctctctacat	cgaaacggga	3900
	atgcagctcg	ccgaacaatt	caaacagaga	gcaatcggtt	tgtctgaaac	agccaccaag	3960
	caaacggagg	ctgctgtctcc	cgtgtgtgaa	tccaagtggc	ggaccctcga	agcctctggy	4020
	gcgaagatca	tgtggaaatt	ctcagcgggg	ttacagagctt	ctgcactctg	agcctctctg	4080
	ctcggaaccc	cccgatagc	atcactgatg	gcattcacag	ccctatcac	gagccctctg	4140
10	accaccaaac	atacctctct	gttttaacat	ctggggggat	gggtggcgcc	ccacttgtct	4200
	ctcccccagc	ctgcttctct	ttctgtaggc	gcgcgcctcg	ctggagccgc	tgttgccagc	4260
	ataggccttg	ggaaggtgct	tgtgatattt	ttggcaggtt	ttggagcagg	gggtgcaggc	4320
	gcgctcgttg	cttttaaggt	catgacggcg	gagatgcctt	ccaccgagga	cctgggttaac	4380
	ctactccctg	ctatcctctc	ccctggcgcc	ctagtctctg	gggtcgtgtg	cgcagcgata	4440
	ctcgtctcgg	acgtggggcc	aggggagggg	gctgtgcagt	ggatgaaccg	gctgatagcg	4500
	ttcgtctcgg	ggggtaacca	cgtctccccc	accactatg	tgtctgagag	cgacgtctga	4560
15	gcagctgtca	ctcagatcct	ctcagtcttt	acactcactc	agctcgtcaa	gaggtctcac	4620
	cagtggatca	acyagagact	ctccacgcca	tgtctcgcct	cgtgtcgtaag	agatgtttgg	4680
	gattggatat	gcacggtggt	gactgatctc	aaagactggc	tccagtccaa	gctcctgcgc	4740
	cgattgcggg	gagtcacctt	ctctctcatg	caacgtgggt	acaaggaggt	ctggcggggc	4800
	gacggcatca	tgcnaaacca	ctgccactgt	ggagcacaga	tccccggaca	tgtgaaaaac	4860
20	ggttcocatg	ggatcgtggg	gcctaggaacc	tgtagttaaa	cgtgcgatgg	aaacttcccc	4920
	ataacggctg	acacacggg	ccctctgacg	ccctcccggg	cgccaaattg	ttctgacggc	4980
	tgtgtggcggt	tgtgtgtctg	ggagtcactg	ggagttacag	ggtgtgggga	tttccactac	5040
	tgtgacgggca	tgacacactga	caacgtaaag	tgcctcgtct	aggttccggc	ccccgaattc	5100
	ttcacagaag	tgagatgggt	gcggtgtcac	aggtacgtct	cagcgtgcaa	acccttccca	5160
	cggggagagg	tcaatctctt	ggctgggctc	aatcaatcac	tggttgggtc	acagctccca	5220
	tgcgagcccg	aaacggagct	acagcgtctc	acttccatgc	tccacgaccc	ctcccaactt	5280
25	acggcggaga	cgctcaagcg	taggtctggc	aggggatctc	ccccctctt	ggccagctca	5340
	tcaagttagcc	agctgtctgc	gcctctcttg	aagggaacat	gcactaccgc	tgtatgacct	5400
	ccggagcgtg	acctcatcga	ggcacaacct	cttggggggc	agagatggg	cggggaacct	5460
	acggcgcgtg	agtcagaaaa	taaggtagta	attttggact	ctttcgagcc	gctccaaagc	5520
	gagggagagt	agaggggaag	actcgtctcg	gcggagatcc	tgcgaggttc	cagggaattc	5580
	ccctcgagca	tgcacctatg	ggcacgcggc	gattacaacc	ctccactggt	agagtcctgg	5640
	aaggaccggg	actacgtccc	tcaagtggta	caacggtgtc	cattgcgcgc	tgcacaaggc	5700
30	ccctccgatac	caactccca	gaggaagagg	acggtgtctc	tgtcaagaat	taccgtgtct	5760
	ttctgcttgg	cggagctctg	cacaaagacc	ttcgagcagt	ccgaactcgt	ggcctctgac	5820
	agcggcagcg	caacggcttc	tcttgacagc	cccttcgacg	acggcgagcc	gggactcgac	5880
	gtgtagctgt	actctcccat	gccccccctt	gagggggagc	ccgggggatcc	cgatctcagc	5940
	gacggctctt	ggctctacgt	aaacggagag	gctagtggag	acgtgtctgt	ctgtctcgat	6000
	tctctacacat	ggacaggcgc	cctgtacacg	ccatgcgctg	cgaggaagaa	caagctgccc	6060
	atcaatgcac	tgagcaactc	tttgtctcgt	caccacaact	tgtgtatcgc	tacaacatct	6120
35	cgacgcgcaa	gctctggcca	gaagaaagtc	acctttgaca	gactgcaggt	cctggacgac	6180
	cactaccggg	acgtgctcaa	ggagatgaag	gcgaaggcct	ccacagttaa	ggcttaactct	6240
	ctatctcgtg	aggaagcctg	taagctgacg	ccccacactt	cgccagatct	taattttggc	6300
	tattggggcaa	agagctctcg	gaacatccac	agaaagcgcg	ttcaacacat	cctctcgtgt	6360
	tggaaaggact	tgtctggaag	caactgagaca	ccaaatgaca	ccaacatcat	ggcaaaaatc	6420
	gaggttttct	gcgtccaaac	agagaagggg	ggccgcgaag	cagctcgctc	tatcgtatct	6480
40	ccagattttg	gggttcgtgt	gttcgcagaaa	atggcccttt	acagatgggt	ctccaccttc	6540
	ccctcagggcg	tgatggggct	ttcatcacga	ttccaatact	ctctctgga	cggggttcag	6600
	ttcctgggtga	atgcctggaa	agcgaagaaa	tgcctatagg	gcttcgcata	tgacacccgc	6660
	tgttttgact	caacggtcac	tgagaatgac	acctcgtttg	aggaatcaat	ctccaatgtt	6720
	tgtgactctg	cccccgaaag	agaaacagcc	ataggtctgc	ccacagagcg	gctgactctg	6780
	ggggggccccc	tgactcaatt	taaaagagag	actcgtcgct	attcgcgtgt	ccgcgcagag	6840
	gtgttactga	cgacacagct	cgttaataac	ctccactggt	acttgaagcg	cgtcgcggcg	6900
45	tgtcgaagct	cgaaagctcca	ggactgcacg	atgctcgtat	gcggagacga	ccttgcgtgt	6960
	atctgtgaaa	gcgcgggggac	ccaaagggag	gaggcggacc	tacgggcctt	cacggagggc	7020
	atgactagat	actctgcccc	cctctggggc	ccgcaccaac	cagaatacga	cttggagtgt	7080
	ataacatcat	gctctccaaa	tgttctcagt	gcgcacgatg	cattgtgcaa	aagggtgtac	7140
	tatctcaacc	gtgacccccc	cacccccttt	gcgcgggctg	cgctggagac	agctagacac	7200
	acttccacga	attctctggt	agccaacatc	actatgtatg	cgccacactt	gtgggcaagg	7260
50	atgctctcga	tgactcaatt	ctcttccact	cttctagctc	agggacacat	tgaaaagcgc	7320
	ctagattgtc	agactcaagg	ggcctgtttac	tccattgacg	actctgacct	acttccagat	7380
	attcaaacag	tccatggcct	tagcgcattt	tccatccata	gttactcttc	aggtgagatc	7440
	aatagggttg	cttcatgcct	caggaaactt	gggtgacgcg	ccttcgcagt	ctggagacat	7500

EP 1 043 399 A2

cgggccagaa gtgtccgcgc taggctactg tcccaggggg ggagggtctc cacttgtggc 7560
 aagtacctct tcaactgggc agtaaggacc aagctcaaac tcaactcaat ccgggtctgcg 7620
 tcccagttgg atttatccag ctgggttcgtt gctgggtaca gcgggggaga catatctac 7680
 5 agcctgtctc gtgccgcacc ccgctgggtc atgtggtgcc tactctact ttctgtaggg 7740
 gtaggcactc atctactccc caacgatga acggggagct aaacactcca ggccaatagg 7800
 ccatcctgtt tttttccctt ttttttttcc tttttttttt tttttttttt tttttttttt 7860
 ttttctcctt tttttttcct ctttttttcc ttttctttcc tttgggtggc ccatcttagc 7920
 cctagtcaag gctagctgtg aaaggtccgt gagcgccttg actgcagaga gtgctgatac 7980
 10 tggcctctct gcagatcaag t 8001

15

20

25

30

35

40

45

50

55

<210> 4
 <211> 7989
 <212> DNA
 <213> Hepatitis C Virus

5

<400> 4
 gccagcccgc gattgggggc gaaactccac catagatcac tcccctgtga ggaactactg 60
 tcttcaagca gaaagcgctc agcactggcg ttatgatgag tgtcgtgcag cctccaggagc 120
 ccccctccc ggaagagcca tagtggtctg cggaaacggf gagtcaacgc gaattgcacg 180
 gagcagcggg tcttttcttg gatcaacggc ctcaatcgct ggagatcttg gcgtgcgcc 240
 gcgaagctgc tagccagtag gtgttggttc gcgaaaggcc ttgtgtgtae gactgatagg 300
 gtgcttgcca gtgcccgggg aggtctcgta gccctgtgca catagacacg aatcctaaac 360
 ctcaaaagaa aaccaaaagg ccgcactgta ttgaacaaag tggattgca cgcaggtcttc 420
 cggccgcttg gtgtggagagg ctattcggtc atgactgggc acaacagaca atcggtcgct 480
 ctgactgcgc cgtgttccgg ctgtcagcgc aggggcgcgc ggttctttt gtcaagacgc 540
 acctgtccgg tgcctgaat gaaactcagg acgaggcagc gcggtatcg tggctgacca 600
 cgcggggcat tcttgcgca cgtgtgctcg acgttgtcac tgaagcgccg agggcctggc 660
 tgcattaggg cgaagtgcgc gggcagatc tctgtctcgc tcaectgtgt cctgcgcaga 720
 aagtatccat catggctgat gcaatggcgc ggctgcatac gcttgatccg gtaactcgc 780
 cattgacca ccaagcgaaa calgcctcg agcgagcacg tactcgatg gaagccgctc 840
 ttgtcgatac ggatgatctg gacgaagagc atcaggggct ccgcgccagc gaactgttgc 900
 ccaggctcaa ggccgcgatg cccgacggcg aggaatcgt cgtgacccat ggcatgctc 960
 gcttgcgcaa tatcatggtg gaaatggccc gctttctcgt attcatcgac tgtggccgcg 1020
 tgggtgtggc ggaacgctat caggacatag cgttggtcac cctgatatc tctgaaagac 1080
 ttgcccggga atgggtcgac cgtgtgctcg gcttttaagg tcttgcgct ccagatcgc 1140
 agcgacacgc ctctctatcg ctctctgagg agttctctcg agtttaacaa gaccacaacg 1200
 gtttccctct agcgggatac attccgcctc tctccctccc ccccccctaa cgttactcgc 1260
 cgaagccgct tggataaagg ccggtgtgctg ttgtctata tgttatcttc caccatattg 1320
 cctgtttttg gcaatgtgag ggcccggaaa cctggccctg tcttcttgac gagoattcct 1380
 aggggtcttt cccctctcgc caaaggaatg caaggtctgt tgaatgtct gaagggaaga 1440
 gttctctcgg aagcttcttg aagaacaaca acgtctgtag cgacctcttg caggcagcgc 1500
 aaccccacac ctggcgacag gtgcctctgc ggccaaagc cacgtgtata agatacaact 1560
 gcaaaaggcg caacaacca gtccacgctt gcgagttgga tagtgtgga aagatcgaaa 1620
 tggctctcct caagctatct caacaagggg ctgaagagtg ccagaagaat acccaattgt 1680
 atgggatctg atctggggcc tcggtgcaac tggttacat gtgtttagtc gaggttaaaa 1740
 aagctctagg cccccegaac cagcgggagc tggtttctc ttgaaaaaca cgataaactc 1800
 atggcgctta ttagcgctta ctcccaacag acgcagagcc taacttgctc catcatcaat 1860
 agcotccacg gccgggacag gaacacaggtc gagggggagg tccaagtgtt tccacccgca 1920
 acacaactct tctggcgac ctgcgtcaat ggctgtggtt ggactgtcta tcatggtgac 1980
 ggctcaaaag ccttgcgcgg cccaaggggc ccaatcaccc aaatgtacac caafgtggagc 2040
 caggacctcg tgggtggca agcgccccc ggggcgcgtt cctgaacac atgcaactgc 2100
 ggcagctcgg accttaact ggtcacagag catgcgagtg tacttcctcg ggcgcggcgg 2160
 ggcgacagca gggggagcct acctccccc aggcocgtct ctaactgtaa gggctcttcg 2220
 ggcggtccac tgcctcgcct ctccgggacg gctgtgggca tctttccggc tgcocgtgtc 2280
 accagagggg ttgcgaagcg ggtgagcttt gtaccocgtg agtctatgca aaccaatg 2340
 cggctcccgg tcttcaacga caactcgctc cctccggcgc taccgcagac attccagggt 2400
 gcccatctac agccocctac tggtagggcg aagaacacta aggtgcgcgc tggatgca 2460
 gcccaagggt ataagggtct tgtcctgaac ccgtccgtcg ccgccacact aggtttccgg 2520
 cgttatatgt ctaaggcaac ttgatctgac ctaacacta gaaccggggt aaggacacac 2580
 accaaggggt ccccaacac gtaactcaac ctatgtaagt tctctgcga cgtgtgtg 2640
 tclggggcgg cctatgaact cataatatgt gatagtgac actcaactga ctgcacactc 2700
 atctcggcca tgcgcacagt cctgacaaac gggagagcgc ctggagcgcg actcgtctgc 2760
 ctccgccacg ctacgctcctc ggatcggttc acogtccac atocaaacat cgaggagggt 2820
 gctcttcca gcaactggaga aatccctctt tatggcgaag ccactcccat cgagaccact 2880
 aaggggggga ggcacactcat ttcttgccat tccaaaga gaatgtgatga gctcccgcg 2940
 aagctgtccg gctcgggact caatcgtgta gcatattacc ggggccttga tgtatcgtc 3000
 atgcaacta gggagagctc tatgtctgta gaaaggagc ctctaatgac ggggtttaac 3060
 gaccattctg actcaagctc agactgcaat acatgtgtca gaagacagct cactctcagc 3120
 ctggacacga cctcaacact caggtgcac acogtgcac aagacgcgtt gtcaactcgt 3180
 cagcggcgag gcaggactgg tagggcgagc atggcgctt acaggtttgt gactcagga 3240
 gaacggccct cgggcattgt agattcctcg gttctgtcg atgtctatga cgcggcgtct 3300
 gcttggtcac agctcaacgc ccgcgagacg tcaagttagt tgcgggtcta cctcaacaca 3360
 ccaggcttgc cagtctgcga ggcacactc g agattctgg agagcgtct taccaggctc 3420
 accacatgat accgcaattt cttgtccagc actaaggac ctctccctac 3480
 ctgtagcat accaggctac ggtgtgcgc agggctcagg ctccacctcc actgtgggac 3540

55

	caaatgtgga	agtgtctcat	acggctaaag	cttaacgtgc	acggggccaac	gcccctgctg	3600
	tataagctga	gagcaattca	aaacgaattt	actcaacac	accocataac	caatacatc	3660
	atggctgga	tgtcgcgctga	ctcggaaggtc	gtccacagga	cttaggtgtat	ggttagcgga	3720
5	ttctcatgca	ctctggcgcc	gtattgectg	acaacagga	gogtggctac	tgtggcgag	3780
	atcattctgt	ccggaagaac	ggccatcatt	cccgacaggg	aagtccttta	ccggaggttc	3840
	gatgaagatg	aagagtcggc	ctcacacctc	cottacatcg	acaagggaat	cagctcgcc	3900
	gaacaattca	aacagaagaag	aatcggtgtg	ctcgaaacag	ccaccaagca	agcgagggtc	3960
	gctgtccctg	tgtgggaatc	caagtgccgg	acctcgaga	cottctcgcc	gaagcatatg	4020
	tgaattctca	tcaaggggat	acaaatatta	gacggcttgc	ctactctgac	tggaacccc	4080
	gogatagcat	cactcgtgcg	attcaacagc	tcttcaacaa	gcocgtctgc	caaccaaat	4140
10	acctctcgt	ttcaactctt	gggggagtgg	gtggccgcgc	aaattgtccc	teccagcctt	4200
	gctcttcgtt	tctgtagggc	cggcatcgct	ggagcggctg	tggcagcatc	agcccttggg	4260
	aagtgctgtg	tgatattttt	ggcaggttat	ggagcagggg	gtggcagcgc	gctcgtggcc	4320
	tttaaggtca	tgagcggcga	gatgccctcc	accgaggaac	tggttaaact	actccctgct	4380
	atcctctccc	ctggcgccct	agtgtccggg	gtcgtgtgcg	cagogatact	gcgtoggcac	4440
	gtgggcccag	gggaaggggg	tgtgcagtgg	atgaaacggc	tgatagcgtt	gcgttcggcg	4500
	gtgtaaacag	tctcccccac	gcaatattgt	cttgagagcg	acgtgtcagc	acgtgtccat	4560
15	cagatctctc	ctagtcttac	catcctcag	ctgctgaaga	ggcttcacca	gtgatacaac	4620
	ggaggaactg	ccacgcccag	ctccggctcg	tggctaaagg	atgtttggga	tgggatatgc	4680
	acgggtttga	ctgatttcaa	gaacctggctc	cagtcocagg	tccgtccgcg	attgcgggga	4740
	gtccctctct	tctcatgtca	acgtgggtac	aagggaactc	ggcgggcgca	cggcatcatg	4800
	caaaaccact	gcccatttgg	agcacagatc	accggacatg	tgaanaacgg	ttccatgagg	4860
20	atcgtggggc	ctagaacctg	tagtaacacg	tggcatggaa	cattccocat	taacgcgtac	4920
	ctcagcggcc	ctcagcggcc	ctcccgcgcg	ccaabattct	ctaggcgctg	gtggcggttg	4980
	gct-gct-gagg	agtaactggc	gggtacgggtg	gtgggggagt	tccctctacgt	gacgggcatc	5040
	accaatgaca	acgtaaagtg	cccggtgcag	gttccggccc	cogaattctt	caacagaagt	5100
	gatgggggtc	ggttgcaacg	gtacgtccca	gogtgcacac	ccctctacag	ggaggaggtc	5160
	accttcctgg	tcgggctcaa	tcaatacctg	gttgggtcac	agctcccatg	cgagcccgaa	5220
	ccggaagtag	cagtgctcac	ttccatgctc	accgaccccc	cccaattacc	ggcgagagcg	5280
25	gctaaagcta	ggctggccca	gggatctccc	ccctccctgg	cagcatctac	agctagccag	5340
	ctgtgtcgcc	cttctctgaa	ggcaacatgc	actaccgctc	atgactcccc	gagctggagc	5400
	ctcactcagg	ccaacctctg	gtggcgagag	gagatagggt	ggaacatcac	cccggtgag	5460
	tcagaanaata	aggtagtaat	tttggactct	ttcgagcgcc	tccaagagca	ggaagatgag	5520
	agggaaagat	cogtctccgc	ggagatccctg	cggaggtcca	ggaaattccc	tcgagcgatg	5580
	cccatattgg	acgcgccgga	ttacaacctc	ccactgttag	agtcctggaa	ggaccgggac	5640
	tacgtccctc	cagtggtaca	cgggtgtcca	ttgcgcctcg	ccaaagcccc	tcgataacca	5700
30	ctctcaacga	ggaaaggagc	ggttgtctcc	tcagaactca	cogtgtcttc	tgccctggcg	5760
	gagctgccca	caaaagacct	cggcagctcc	gaatcgctcg	cogtgcagcg	cggcagcgca	5820
	acggcctctc	ctcagcagcc	ctccgcagac	ggcgacgcgg	gatcgacgtg	tgagctgcac	5880
	tccctccatgc	cccccttgga	gggggagctg	gggagctccg	atctcaagga	cgggtctcgt	5940
	tctaacgtaa	gcggaaggcg	tagtaggagc	gtcgtctgct	gctcgatgc	ctacacatgg	6000
	acagcgccccc	tgatcacgcc	atgcgtctgcg	gaggaaacca	agctgcocat	caatgcactg	6060
	agcaactctt	tgtctcgtca	ccacaacttg	gtctatgcta	caacatctcg	cagcgcaagc	6120
35	ctcgcgacga	agaaagtca	ctttgacaga	ctcgaggtcc	tggacagcca	ctaccgggac	6180
	gtgctcaagg	agatgaagcg	gaagcgctcc	acagttaagg	ctaaactctt	atccgtggag	6240
	gaagcctgta	agtcagcgcc	ccccaatcgc	ggcaatcata	aatttggcta	tggggcaag	6300
	gtcgttcgga	acctatcgac	caaaggcgtt	atccacatcc	gctccgtgtg	gaaggaattg	6360
	ctgggaagaca	ctgagacacc	aattgacacc	accatcatcg	caaaaatgga	ggttttctgc	6420
	gtcccaacga	agaaaggggg	cogcaagcca	gctgcoccta	togtattccc	agatttgggg	6480
40	gttcgtgtgt	gcgagaanaa	ggccctttac	gatgtggtct	caaccctccc	ctaggccgtg	6540
	atgggctctt	catacgattt	ccaataactc	cttggaacgc	gggtcgagtt	ctgtgtgaa	6600
	gctcggaaag	cgaagaanaat	cccatatggc	ttcgcatatg	caaccocgct	ttttgactca	6660
	acggtcacctg	agaaatgcat	ccgtgttgag	gagtcacatc	accaatgttg	tgacttggcc	6720
	ccgaagcgga	gagagcccat	aaggtcgctc	acagacggcg	tttaacagg	gggccccctg	6780
	actaatctta	aagggaagaa	ctcgtgctat	cgccggtctc	gacagagggg	tgacttgaac	6840
45	acacagctcg	gtaataacct	caatgtttac	tgaagcgccg	ctcgccgctg	tcagagctcg	6900
	aagctccagg	actgcacagt	gctcgtatgc	ggagacgacc	tttgtgttat	ctgtgaaagc	6960
	cggggagacc	aagaggacga	ggcgagccta	cgggccttca	cggaggtcat	gactatgata	7020
	tctgcgccccc	ctggggagccc	gcccacaacca	gaatacga	ctggagtgtg	acaatcatgc	7080
	tctctcaaat	tgtcaagtgc	gcacagatga	tctgcacaaa	gggtgtacta	cttccacoggt	7140
	gaccocacaa	ccccctctgc	cgccgctcg	tgggagacag	ctagacacac	tcacagtcaat	7200
	tccgtgctag	gcaacatctg	ctgtatgcg	ccccactctg	gggcaagagt	gactccgtatg	7260
50	actcaattct	tctcactctc	tctagctgat	gaaacaactg	aaaaagccct	agatttggca	7320
	atctcagggg	cctgttactc	cattgacaga	cttgcactac	ctcagatcat	tcaacagctc	7380
	catgccccta	gcgcatttcc	actccatagc	tactctccag	gtgatatcaa	taggggtggc	7440
	tcatgcctca	ggaaacttgg	ggtaccgcgt	ttgogagtct	ggagacatcg	ggccagaagt	7500

EP 1 043 399 A2

gtccgcgcta ggctactgtc ccaggggggg agggctgcc a ttgtggcaa gtacctcttc 7560
aactgggcag taaggaccaa gctcaaatc actccaatcc cggctgcgtc ccagttagat 7620
5 ttatccagct gggtcgttgc tggttacagc gggggagaca tatatcacag cctgtctcgt 7680
gcccgacccc gctggttcat gtggtgccta ctccacttt ctgtaggggt aggcacatat 7740
ctactcccca accgatgaac ggggagctaa acactccagg ccaataggcc atcctgtttt 7800
tttccctttt tttttttctt tttttttttt tttttttttt tttttttttt ttctcctttt 7860
tttttcctct ttttttcctt ttctttcctt tggtagctcc atcttagccc tagtcacggc 7920
tagctgtgaa aggtccgtga gccgcttgac tgcaagagat gctgatactg gcctctctgc 7980
10 agatcaagt 7989

15

20

25

30

35

40

45

50

55

<210> 5
 <211> 8649
 <212> DNA
 <213> Hepatitis C Virus

5

<400> 5
 gccagccccc gattgggggc gacactccac catagatcac tcccctgtga ggaactactg 60
 tcttcaacga gaaagcgtct agccataggc tagtatgagc tgcgtgtcag cctccaggac 120
 ccccccctcc ggaagagcca tagggctgtg cggaaacggg gagtacaccc gaattgtcac 180
 gacgacccgg tctttctgt gatcaacccg ctcaatggcc ggaagattgg gcgtgcccc 240
 gcgagactgc tagccgagta gtgttggggt gcgaaaggcc ttgtgtact gccctgtagg 300
 gtgcttgcga gtgcgccggg aggtctcgtg gccctgtcac catgagacgc aatcctaacc 360
 ctcaagaata aaccaaacgt aacaccaacg gcgcgcccat gattgaacaa gatggattgc 420
 acgcagggtc tccgcgcqct tgggttgaga ggctatttcg ctatgactgg gcacacaga 480
 caatcgcgct ctctgatgcc gcgctgttcc ggctgtcagc gcaggggcgc ccggttcttt 540
 ttgtcaaacg cgactcgtcc gggtgccctga atgaactgca ggaacaggca gcgcgcctat 600
 cgtggctggc cagcaggggc gtctccttgc cagctgtgct gaacgttgtc actgaagcgg 660
 gaagggactg gctgtatgtt ggcgaagtcg ccggttgaga tctctgtca tctcaccttg 720
 ctcttcacga gaaagattcc atcatagctc atgcaatgcg gcgcgtgcac acgcttgatc 780
 cgcgtacact gccattogac caccaaagca aacatgcgat cgaagagca cgtactccga 840
 tggaaagccg tctgtcgat caggatgata tggacgaaga gcaatcaggg ctgcgcgcag 900
 ccgaactgtt cgcagcgctc aaggcgcgca tgcccagcgy caggagatct ctgctgaccc 960
 atggcgatgc ctgcttgccc aatatcatgg tggaaaatgg ccgctttctt ggaattatcg 1020
 actgtggccg gctgggtgtg gcggaacgct atcaggacat acgcttgctt accgtgata 1080
 ttgctgaaga gcttggcggc gaattgggctc acgccttgct agtgccttac agtatcgccg 1140
 ctcccgatcc gacgcgcatc cccttctatc gccttcttga cgtctcttga tgaattttaa 1200
 cagaccacaa cgttttccct cttagcggatc caattccgcg cctctccctc cccccccctc 1260
 aacgttactg gcgcgaagcc cttggaaata gcccggtgtg cgttttgtca tatgttattt 1320
 tccaccatat tgcgctcttt tggcaattgt agggcccoga aacctggccc tgtctctgtc 1380
 agcagcattc cttaggggtct tcccctcttc gccaaaggaa tgaacggtct gttgaattgc 1440
 tgaaggaagc cagtctcctt ggaagctctt tgaagacaa caacgctctg agcacacctt 1500
 taagggcagc ggaacccccc acctggcgac aggtgcctct gcggccaaaa gccacgtgta 1560
 tgaatatac ctgcgaagcc ggcaaacacc cagtgcacgc tgytggattg tagatgtgtg 1620
 gaaagagcca aatgtctctc ctcaagcgtt tcaacaaagg ggtctgaagg tgcacgaagg 1680
 ctgccccatt gtatgggaat tgatctgggg cctcggtgca catgctttac atgtgtttag 1740
 tccaggttaa aaaaagctta gcccccoga accacgggga cgtgtgtttc ctttgaataa 1800
 caccgataata ccatggaccg ggagatggca gcatcgtgtg gaaggcgcggt ttctgtaggt 1860
 ctgataactc tgaacctgtc accgcaatat aagctgttcc tgcctaggct catatgggtg 1920
 ttacaatttt ttataccagc ggccgaggca cacttgcaag ttgtgattcc cccctcaac 1980
 tctcgggggg gcgcgcgctc cgtcatctct ctacgctgc cgtatccacc agagtatac 2040
 ttatacatca ccaaaattct gctgcgcata ctcggtccac tcatgtgtct ccaggtaggt 2100
 ataaaccaag tgcgctactc cgtgcgcgca cagggctcca tctgtgcatt catgctgtgt 2160
 cgaaagtttg ctggagatca ttatgtccaa atagcttcca tgaattggc gcactatgca 2220
 ggtaagtacg ttatgatcca tctcacccca ctgcgggact ggcccacagc ggccctaaga 2280
 gacacttcgg tggcagttga gcccgctgtc ttctctgata tggagaccaa ggttatcacc 2340
 tggggggcag acacccggcg gtgtggggac atcatcttgg gctgcgccgt ctccggccgc 2400
 agggggaggg agatacatct ggccccgcca gacagccttg aaggcgagg gtgcgaactc 2460
 ctccgcctta ttccggccta ctcccacaga acgcgaggcc taacttgctg catcataact 2520
 agctctcaag gcgggaagac gaacaaagtc gggggggagg tccaagtgtt ctccacgta 2580
 acaactctt tctggcgac ctgctcgaat ggcggtttgt ggaactgta tcatgtgga 2640
 ggctcaaga cctctggcgg cccaaaaggc ccacacccc aagtatcac caatgtgac 2700
 caggacctcg tccgttgga ggcgcgcccc ggggcgcgtt ccttgacacc atgcacctg 2760
 ggacgtccgg acctttactt ggtaacagg catgcgatt tcaattccgt gcgcggcgcg 2820
 ggccacagca gggggaggct actctcccc accgcctctg ctacttgaa gggctctctg 2880
 ggcggtccac tgcctgcgcc ctccgggcca cgtgtgggca tctttcgggc tgcgtgtg 2940
 acccgagggg ttgcgaagcg ggtggaattt gtaccocgtg agtctatgca aaccaatg 3000
 cgttccccgg tcttcaagca caactcgtcc cctccggcgc taccgcagac attccaggtg 3060
 gccactcac accgcctcac tggtagcgcc aagaacaccc aggtgcgcgc tgcgtatgca 3120
 gccaaagggt ataagttgct tctctgaac ccgtccgtcc ccgccaacct aggttctcgg 3180
 gcgtatagct ctaaggacaa tggatccagc cctaactcca gaacccgggt aagaccatc 3240
 accacggctg cccccaacac gtactccacc tatggcaagt ttcttgccga cgtgtgtg 3300
 tctggggggc cctatgacat cataatatgt gatgagtccc actcaactga ctgcacact 3360
 attctgggca tggcagcagt cctgacacaa gcgagagacg ctagagccgc actcgtgtg 3420
 ctccgcacgy ctaagcctcc gggtctcgtt acgctgccac actcaaacct cagaggaggt 3480
 gctctgtcca gcaactgaga aatccctttt tatggcaagg ccatcccatc cgagaccatc 3540

55

aaggggggga ggcacctcat tttctgcaat tccaagaaga aatgtgatga gctgcgcgg 3600
 aagcgtgcgg gctctggact caatgctga gcatattcac gggcgttga tgatattcac 3660
 ataccacta cgggagcgt cattgtctga gcaaggacgt ctctaatacg ggcctttacc 3720
 ggcgatttgc actcagtgat ggaactgcaat acatgtgtca cccagacagt cgaactcagc 3780
 ctggaccocga ccttcacacat tgagacgacg accgtgcaac aagacgcggt gtcaacgtcg 3840
 cagcgcggag cgaagcactgg tagggccagg atgggcattt acaggtttgt taactocaga 3900
 gaacgcgcct cgggcatggt cgattctcgt gtctgttgcg agtgcattga cgcggcgtgt 3960
 gctttggtacg agctcaagcc cgcgcgagcc tcaagttagt tggcggtcta cctaaacaca 4020
 ccagggttgc gactgtgaca ggaactctgt gactgttgg agaggtctct taacaggctc 4080
 accccaatga agcccaattt ctgttccagc actaagcagg caggagacaa ctccctctac 4140
 ctggtagcat accaggtctac ggtgtgcgcc aggtctcagg ctccacctcc atcgtgggac 4200
 caaatgtgga aqgtgtctat accgttaaag cctacgtctg acgggcacac gccctctgct 4260
 tatagctggg gaggcgttca aaacgaggtt actaccacac accccataac caaatatcat 4320
 atggcatgca tgtcgcgtga cctggaggctc gtcacagaca cctgggtgct ggtaggcgga 4380
 gctctagcag ctctggcgcc gattgtcgtg acaacagcca gctgtgctat tctgtggcag 4440
 atcaccttgt ccggaagacc ggcactcatt ccgcagaggg aagtccttta ccggaggttc 4500
 gatggatgg aagagtcgct ctccacacct ccttaacctc aacagagcat cgcagctggc 4560
 gancattcca aacagaggc aatcgggttg ctgcaaacag ccaaccaaga agcgaggctc 4620
 gctgctcccg tgggtgaatc caagtggcgg accctggaag cctcttgggc gaagcatatg 4680
 tggaaattca tcagcgggat acaatattta gcaaggttgt ccaactgtgc tggcaacccc 4740
 gcgatagcat cactgatggc attcacagcc tctatcacca gcccgctcac caccacaact 4800
 accctctcgt ttaacatcct ggggggatgg gtggcgcccc aacttgctcc tccagcgtct 4860
 gctcttgctt tctgtaggcg cggcatctgt gggagggcgt ttggcagcat aggccttggg 4920
 aagcgttga tgaattattg ggaactctat gggcaggggg tggcgaggcg tgcgtggccg 4980
 ttaaaagtca tgagcggcga gttgcctctc accgagagct tggcttaact atcctctgct 5040
 atcctctccc ctggcgccct agtctctcgg gtctgtgtgc cagcgtactc gctgtggcac 5100
 gtgcggccag gggagggggg tgtgcagttg atgaacggcg tgatagcgtt cgtctggcgg 5160
 ggttaaccac tctccccac gcaactatgt cctgagagcg accgtgcagc acgtgtcact 5220
 cagatctctc ctatgtctac ctactctcag ctgtcgaaga gcttccacca gttgactcaac 5280
 gaggactgct ccaagcccatg ctccgctcgt tggctaaag atgtttggga ttggatgacg 5340
 agcgtgtga tgaatttcaa gaactggctc cagtccaagc tctcgcgcg attgtcggga 5400
 ctccctctct tctcatgtca acgtgggtac aaggagctct ggcggggcga cctgactgtg 5460
 caaacacctc gccatgttgg agcacagatc accggacatg tgaaaaaggc ttccatgagg 5520
 atcgtggggc cttagcacctg tagtaaacag tggcatggaa cttccccatc taacgctgac 5580
 accacggggc cctgcacgcc ctccccggcg ccaatttatt ctaggcgctc gtggcggttg 5640
 gctgctgagg agtacgttga ggttacgcgg gtgggggatt tccactacgt gacgggcatc 5700
 accactgaca acgtaaagtg cccgtgtcag gtccggcccc ccgaattctt cacagaagtg 5760
 gatggggctg ggtttcacag gtaacttcca cgtgtcaaac cctctctacg ggaagagctc 5820
 acattctcgg tgggtgtcaa tcaaaactcg gttggttca agtctccatg cgaagtcgaa 5880
 cgcagcgtag cagtgtctac ttcactgtct accgaacctc cccaattac ggcggagagc 5940
 gctgaagcta ggcgtgcgag gggatctccc cctccttgcc ccagctcacc agctagccag 6000
 ctgtctgcgc cttccttgaa ggcaacatcg actaccogtc atgactcccc ggacgtgacg 6060
 ctcatcgagg ccaacctctc gtggcggcag gagaatggcg ggaacatcac ccgctgggag 6120
 tcagaaaaa aggtagtaat tttgactctc ttcagagccc tcacagcgga gaggatgag 6180
 agggaaatgt cgttccggc ggaagactct cggagatcca ggaattcccc tccagcgatg 6240
 cccatctggg cagcccgga tacaacacct ccaactttag agtctgggaa ggacccggag 6300
 taactccttc caatgttaca cgggtgtcca ttgcgcctg ccaaggcccc tccgatacca 6360
 ccttcacgga ggaagaggac ggtgttctct tcaaatctca cctgtctctc tcccttggcg 6420
 gagctgcgca caaagacctt cgcgacgtcc gaactgtcgg cctgcgcagc cgcacggcga 6480
 accgctctc ctgaccagcc ctccagagac ggcgacgcgg gatccagcgt tgagtgtcac 6540
 tctctcatcg ccccccctga ggggagcgcc ggggatcccg atctcacgga cgggtcttgg 6600
 tctacgttaa cgcagggaggc tagtgaggac gtctctgtct gctcgtatgc ctacacatgg 6660
 acagggcgcc tgatcacgcc atgcgtctgg gaggaaacca agtgcacctc caatgcactg 6720
 agcaactct tgcctcgtct ccaactcttg tctgatgca caactctgct cagcgcagag 6780
 ctgcggcaga agaaggtcac ctttgacaga ctgcaggtcc tgaacagcca ctacccggag 6840
 gctgtcaagg agatgaagg cgaagcgctcc acagttaagg ctaaatctct atccgtggag 6900
 gaagcctgta agctgaogcc ccaaccattg gcaagatcta aatttgctta tggggcgaag 6960
 gacgtccgga acctatccag caaggccggtt aaccacatcc gctccctgtg gaaggacttg 7020
 ctggagaaga ctgagacacc aattgacacc acctatgg gcaaaaatga gttttctgcg 7080
 gtccaacagg agaagggggg ccgaagacca gctgcctcta tctgattccc agatttgggg 7140
 gttctgttgt cgcagaaaaa ggcocctaac gatgtggtct ccaacctccc tggagcgttg 7200
 atggtgctct ctatagattt ccaattgtct ctgacagc ggttcagatt cctagtgtac 7260
 gctcggaaag cgaagaaatg ccttatggcg ttgcattatg acccccgctt ttttgactca 7320
 accgtcactg agaattgcat aaggtctgctc acagagcgcg ttaacatgct tgaactggcc 7380
 ccgaagacca gacaggccat aaggtcgctc acagagcgcg ttaacatgct gggccccctg 7440
 actaatctca aaggcgagaa ctgcgcgtat cgcggtgccc ggcgcagcgg tgaactgacg 7500

accagctgog gtaataccct cacatgttac ttgaaggcog ctgcggcctg tgcagctgog 7560
 aagctccagg actgcacgat gctcgtatgc ggagacgacc ttgtcgttat ctgtgaaagc 7620
 gcggggaccc aagaggacga ggcgagccta cgggccttca cggagggctat gactagatac 7680
 5 tctgcccccc ctgggggaccc gcccaaacca gaatacgact tggagttgat aacatcatgc 7740
 tctcccaatg tgtcagtcgc gcacgatgca tctggcaaaa ggggtgtact tctcaccogt 7800
 gaccccaacca ccccccctgc gcgggctgog tgggagacag ctgacacac tccagtcaat 7860
 tctcggctag gcaacatcat catgtatgog cccaccttgt gggcaaggat gatcctgatg 7920
 actcatttct tctccatcct tctagctcag gaacaacttg aaaaagccct agattgtcag 7980
 10 atctacgggg cctgttactc cattgagcca cttagacac ctcagatcat tcaacgactc 8040
 catggcctta gcgcatttct actccatagt tactctccag gtgagatcaa taggggtggct 8100
 tcatgcctca ggaaacttgg ggtaccgccc ttgcgagtct ggagacatcg ggccagaagt 8160
 gtccgcgcta ggctactgtc ccaggggggg agggctgcca ctgtggcaca gtacctcttc 8220
 aactgggcag taaggaccaa gctcaaacct actccaatcc cggctgcgtc ccagttggat 8280
 ttatccagct ggttcgttgc tggttacagc gggggagaca tatacacag cctgtctcgt 8340
 15 gcccgacccc gctgtttcat gtggtgccta ctactacttt ctgtaggggg aggcattctat 8400
 ctaetcccca accgatgaac ggggagctaa acactccagg ccaataggcc atcctgtttt 8460
 ttcccttttt ttttttcttt tttttttttt tttttttttt tttttttttt ttctcctttt 8520
 tttttctctt ttttttcttt ttctttcttt tgggtggctcc atcttagccc tagtcacggc 8580
 tagctgtgaa aggtccgtga gccgcttgac tgcagagagt gctgatactg gcctctctgc 8640
 20 agatcaagt 8649

25

30

35

40

45

50

55

<210> 6
 <211> 8001
 <212> DNA
 <213> Hepatitis C Virus

5

<400> 6
 gccagccccc gattgggggc gaactccac catagatcac tccctgtgga ggaactactg 60
 tcttcaacga gaaagcgctc agcactggcg ttagtatgag tgtcgtgcag cctccaggag 120
 ccccccctcc gggagagaca tagtggctcg cggaaacggt gagtaacacg gaattgccag 180
 gagccacggg tcttctcttg gatcaacccg ccaatcgctc ggaattttgg ggtggccccc 240
 10 ggcagctcgc tagccagtag gtgttgggtc gogaaagggc ttgtggtaac gcttgatagg 300
 gtctcttgca gtgccccggg aggttctcgt gaccgtgcac catgagcagc aatctctaac 360
 ctcaaaagaa aaccaaacgt aaccaaacgc ggcgcgcact gattgaacaa gatggattgc 420
 acgcaggttc tcggcgccgt tgggtggaga ggtcattcgg ctatgactgg gcacaacaga 480
 caatcgctcg ctctgatgcc gcgtgttccc ggtcgtcagc gcaggggcgc ccggttcttt 540
 15 ttgtcaagac cgactcgtcc ggtgcctcga atgaactgca ggaacaggca gcggtgctat 600
 cytgcgtggc caagacgggc gttccttggc cagctgtgct cgactgtgtc actgaagcgg 660
 gaaggagctg gctgctattg ggcgaagctg cggggagaga tctcctgtca tctcaacttg 720
 ctctcgccga gaaagtatcc atcatgggtg atgcaatgag gggcgtgcac acgtctgtac 780
 ccgtctacct ccacttogac caaccaagca aacatcgcat cgaacagcga cgtactcgga 840
 tggaaagccg tcttctcgat caggtgatcc tggacgaaga gcatcagggg ctccgcccag 900
 ccgaactgtt ccgcaggctc aagcgcgcca tgcccagcgg cgaggatctc gtgctgacc 960
 atggcgatcg ctgcttgccc aatatcatgg tggaaaatgg ccgtttttct gattcatcgc 1020
 20 actgtggcgc gctgggtgtg gcgacgcgct atcaggacat agcgtttgct accgtgata 1080
 ttgctgaaga gcttggcggc gaatggggtc accgtctctc gctgtcttca ggtatcgccc 1140
 ctcccgatgc ctgcgcgcat gccctctatc gccctcttga ccgttctctc aggtctttaa 1200
 cagacaccaaa cggtttccct ctacggagat caattccgc cctctccctc ccccccctcc 1260
 aacgttactg gccgaagcgc cttggaataa ggcgcgtgtg cgtttgtcta tagttatttt 1320
 25 tccacacatc tgcggtcttt tggcaatgtg agggcccgga aacctggccc tgcctctgtg 1380
 acgagcatcc ctacgggtct tccccctcc gccaaaggaa tgcaaggtct gttgaatgc 1440
 gtggaaggag cagttcctct ggaagcttct tgaagacaaa caacgtctgt agcagccct 1500
 tgcagcgagc ggaaccccccc acctggcgac aggtgcctct cggcgccaaaa gccactgta 1560
 taagatacac ctgcaaaagg ggcacaaccc cagtgccacg ttgtagcttg gatagtgtgt 1620
 gaaagagcta aatggctctc ctcaagotta tccaaacagg gctgaagga tgcacagaa 1680
 gtaccacctt gtatgggata tgactgggg cctcgttgca catgctttac atgtgttag 1740
 30 tccaggttaa aaaaagtcta gcccocccga accacgggga cgtggttttc ctttgaaaaa 1800
 cagcataata ccattggcgc tattacggcc tactcccaac aga cggcgga cctacttgcc 1860
 tgcatcatca ctacgctcac aggcgggacg aggaacacag tcgaggggga ggtccaaagt 1920
 gtctccaccc caacacaatc ttctcggcgc acctcgctca atggcgtgtg ttggactgtc 1980
 tatcatggtg ccggtctcaaa gaaccttgcc gaagccacgg gcccaatcac ccaaatgtac 2040
 accaatgtgg accaggacct cgtcgtctgg caagcgcccc ccggggcgcg ttctttgaca 2100
 ccatgcacct gcggcagctc ggaactttac ttggtcaaga ggcattgcga tgcattccg 2160
 35 tgcgcgcgcg gggcgacaag cagggggagc ctactctccc ccaggccgct cctactttg 2220
 aagggctctt cggcggttcc actgctctgc cctccggggc atgctgtggg catctttcgg 2280
 gctgcggtgt gcaacccgag ggttgcaag gcggtggaat ttgtacccgt agactctatg 2340
 gaaacacata tgcggtccccc ggtcttcacg gacaactcgt cccctccgcg cgtaccgcag 2400
 acattccagg tggcccatct acacgcacct actggtagcg gcaagagcac taagggtccc 2460
 gctcgtatgt cagcccaagc gtataaaggt cttgtcctga acccgtccgt ccgcgccacc 2520
 40 ctagggttgt gggcgtatat gcttaaggca catggtatoy accctaaact cagaaacggg 2580
 gtaggagaca taacaacggg tgcccacatc agcactcoca cctatggcaa gttctctgca 2640
 gacgtgtgtt gctctggggg cgcctatgac atcataatct gttgatgagt ccaatcaact 2700
 gactcagaca ctactcgggg catcgccaca gctcctgaa cagcggagac ggtcgtgagc 2760
 cgactcgtcg tgctcgccac cgtatagcct ccggatggat taccgtgccc acatccaaac 2820
 atcgaggagg ttgctctgtc cagcatcgga gaataccctc ttatggcaa agcatccccc 2880
 atcgcgacca tcaagggggg gaggcacctc atttctgccc attccagaaa gaaatgtgat 2940
 45 gactcgcgcg cgaagctatc cgcctccgga ctcaatgctg lagcatatta ccgggacctc 3000
 gatgtatccg tcataccaac tagcagagac gtcaat'gtcg tagcaacgga cgtctaatg 3060
 acgggcttta ccgggatatt gactcagtg atcgactgca atacatgtgt caccacagca 3120
 ctgactctca gctcgaaccc gaactcaacc attgaagaca cagactgccc acaagacgg 3180
 gtgtcacgct cgcagcgggc aggcagagact ggtaggggca ggaatggcat tacaagttt 3240
 gtgactccag gagaacggcc ctccggcatg ttcgattctc cgttctctgt cgaagtctat 3300
 50 gacgcgggct gtcgtttgta cgaactcaag cccgcggaga cctcagtttg gttcgggct 3360
 taccttaaaa cacaagggct gccgctctgc caggacacac tggagttctg ggaagcgtc 3420
 tttaacggcc taaccacatc agacgccact tctctgtccc agactaggga ggcagagact 3480
 aacttccccct acctggtagc ataccaggct acgggtgtgc ccagggtcca gctccacct 3540

55

	ccatgttgagg	accaaagtgtg	ggagtgcttc	atacggctaa	agcctacgct	gcacggggca	3600
	acgcccctgc	tgtatagctg	agagccgttt	aaaaagagag	ttactacac	acaccccta	3660
	accaaataca	tcatggcatg	catgtogctg	gacctggagg	tcgttcaacg	cacctgggtg	3720
5	ctgttaggag	gagtcctagc	agctctggcc	gcgtattggc	tgacaacagc	cagctgtgtc	3780
	tattgtggga	ggatcatcct	gtccggaaag	ccggccatca	ttcccgacag	ggaagtccct	3840
	caacgggagt	togatgagat	ggaagatgac	gectcacacc	ttccctacat	cgaacaggga	3900
	atgcagctcg	cogaacaatt	caaacagaa	gcaatccggg	tgcctgaaac	agccaccaag	3960
	caacgggagg	ctgtctgtcc	cgctggggaa	tccaagtggc	ggacccctga	agcctctctg	4020
	ggagagccta	tgttgatatt	catctggggg	atacaatttc	tacagaggtt	gtccactctg	4080
	ctgtggcaac	ccgggatagc	atacattgag	gcattccacg	ccctctatcc	cagcccgctc	4140
10	accacccaac	ataccctcct	gtttaacate	ctggggggat	gggtggccgc	ccaacttgct	4200
	ctccccagcg	ctgctctctc	tttcttaggc	gccggcatcg	ctggagcggc	tgtttgcacg	4260
	ataggccttg	ggaaggtgct	tgtggatatt	ttggcaggtt	atggagcagg	ggtggcaygc	4320
	gcgctcgtcg	ccctttaaagt	catgagcgcc	gagatgcctt	ccaocaggga	ctgtgttaac	4380
	ctactccctg	ctatcctctc	ccctggcgcc	ctagtctgct	gggtctgtgt	cgacgcgata	4440
	ctgctcggcc	cgtggggccc	aggggagggg	gctgtgcagt	ggatgaaocg	gctgatagcg	4500
	ttcgtctcgc	ggggtaacct	cgtctccccc	acgcactatg	tgtctggagc	cgcagctgca	4560
15	gcacgtgtca	ctcagatcct	ctctagtctt	accaacactc	agctctgtga	gaggtctcac	4620
	cagtggtaca	acqaggaect	ctccacgcga	tgctccggct	cggtgcctaa	agatgtttgg	4680
	gatttgatag	gcacgctggt	gactgatttc	aagacctggc	tccagtccaa	gctcctgcgc	4740
	gcgtctgcgg	gagtcctcct	ctctctcatg	caacgtgggt	acaagggaqt	ctggcggggc	4800
	cagggcatca	tgcaaaacac	ctgcccattg	ggagcacaga	tcacgggaca	tgtgaanaac	4860
20	gggtccataga	ggatctgtgg	gctcagggac	tggttaacaa	cggtggcatgg	aaacttcccc	4920
	attaaocgtg	acacgcgcgc	ccctcccgcg	ccctcccgcg	cgcacaatta	ttctatgggcg	4980
	ctgtggcggg	tggctgtctga	ggagtaactg	bagtttaacg	ggttggggga	tttccactca	5040
	gtgacgggca	tgacacactg	caacgttaag	tgccctgttc	aggttccggc	ccccgaattc	5100
	ttcacaaga	tgagttgggt	gcggttgcac	aggtacgctc	cagcgtgcaa	acccctcta	5160
	cgggaggagg	tcacattcct	ggtcgggctc	aatcaataac	tggttggttc	acagctcccc	5220
	tgcgggcccg	aacccgacgt	agcagtgctc	acttccatgc	tcacgcgacc	ctccccactt	5280
25	acggcgggaga	cggctaaagc	lagctctggc	aggggatctc	cccccctctt	ggccagctca	5340
	tcaagctagc	agctgtctgc	ggcttccctg	aaggcaaatc	gactactccg	ctatgactcc	5400
	cgggagcctg	acctaatcgc	ggcccaacct	ctgtggcgcc	agagataggg	cgggaacatc	5460
	acccgcgtcg	agtcagaaaa	taaggttagta	attttggact	ctttcgagcc	gctccaagcg	5520
	gaggagagat	agagggaagt	atccgttcgc	gcggagatcc	tgccgaggtc	caggaaattc	5580
	ccctcagcga	tgcccatatg	ggcaccgcgc	gattacaacc	ctccactggt	agagtctcgg	5640
	aaggaccocg	actacgtccc	tccagtggta	caocgggtgct	catltgcgcg	tgccaaggcc	5700
30	ccctcagatc	caactccacg	gagggagagg	acggttgtcc	tgtcagaatc	tacgtgtctt	5760
	ctgcctctgg	cggagctctc	cacaaagacc	ttccgagctc	cogaatctgtc	ggccctcgac	5820
	agcgacacgc	caacggcttc	ttctgcacg	ccctccagcg	acggcgagcc	ggagtcagac	5880
	gtgtagtcgt	actcctccat	gcctcccctt	gagggggagc	cgggggatcc	cgatctcaag	5940
	gacccgtctt	ggtctacogt	aagcgaggag	gotagtggg	acgtctctct	ctgtctcgatg	6000
	tcctacacat	ggacaggcgc	ccgtcatcgc	ccatgcgcgt	cgggagaaac	caagctgcgc	6060
	gtcatgcac	tgagcaactc	tttgcctcgt	accaccaact	tggtctatgc	tacaaacatc	6120
35	cgcagcgcga	gcctcggggc	gaagaaggtc	acctttgaca	gactgcaggt	ctcggacgac	6180
	caactccggg	acgtgctcca	ggagatgaag	ggcaaggcgt	ccacagttga	gagctaaact	6240
	ctatcccggt	aggaagctgt	taagctgagc	cccccactat	cgggccagatc	ctaatattgg	6300
	tatggggcaa	agggcgtccg	gaactatccc	agcaaggccg	ttaacacatc	ccgtccctgt	6360
	tggaaggact	tgtctggaag	cactgagaca	ccaattgaca	ccaacctcat	ggcaaaaaat	6420
	gaggttttct	gcgtccaacc	agagaagggg	ggccgcaagc	cagctcgctt	tatcgtattc	6480
40	ccagattttg	gggtttcgtg	gtgogagaaa	atggcccttt	acgagtgtgt	ctccaccttc	6540
	ctccagcgcg	tgatgggctc	ttcatcagga	ttccaatact	ctcctggaca	gcgggtcgag	6600
	ttctcgttga	atgcctggaa	agcgaaagaa	tgccctatgg	gcttgoata	tgaacccgcg	6660
	tggtttgact	caacggtcac	tgagaaatgc	atccgtgtgt	agaggtcaat	ctaccatgtt	6720
	tgtagctgtg	cccccgaagc	cagcaagcgc	ataagtcgcg	tcacagagct	gotttaacat	6780
	gggggccccc	tgactaaatc	taaagagcag	aacttgcgtg	atccgctgct	cgcgcggagc	6840
	gggttactga	cgaccagctg	cgttaataac	ctcacatggt	atttgaaggc	cgtctcgccg	6900
45	tgctcagctg	cgaagctcca	ggacgtcacg	atgctcgtat	gcggagacga	cccttgttgt	6960
	attctgtaaa	gcggcgggga	ccaagaggac	gagggcgagc	tacgggctct	cacggaggct	7020
	atgactagat	actctgcccc	ccctggggac	cgcgccaac	cagaatacga	cttggagtgt	7080
	ataacatcat	gctcctccaa	tgtgtcagtc	gcgcagaatg	catctggcaa	aaggggtgac	7140
	tatctcaccc	gtgaccccca	cacccccttt	gcggggctgt	ctgtggagac	agctagaac	7200
	actccagtcg	attctctgct	agccaagctc	ataagctatg	ctgacagctt	gtgggacagg	7260
50	atgatccctg	tgactcattt	cttctccatc	ctcttagctc	aggaacaact	tgaaaaagcc	7320
	ctagattgtc	agatctacgg	ggcctgtttc	tcacattgac	cacttgacct	acctcaagtc	7380
	attccacgac	tcacatggct	tacgcatttt	tcactccata	gttactctcc	aggtgagatc	7440
	aataggggtg	cttcatgcct	caggaaactt	ggggtaccgc	cttgcgaggt	ctggagacat	7500

EP 1 043 399 A2

5 cgggccagaa gtgtccgcgc taggctactg tcccaggggg ggagggtgc cacttgtggc 7560
 aagtaacctct tcaactgggc agtaaggacc aagctcaaac tcactccaat cccggctgcg 7620
 tcccagttgg atttatccag ctggttcggt gctggttaca gcgggggaga catatatcac 7680
 agcctgtctc gtgcccgacc ccgctggttc atgtggtgcc tactcctact ttctgtaggg 7740
 gtaggcattct atctactccc caaccgatga acggggagct aaacactcca ggccaatagg 7800
 ccattcctgtt tttttccctt ttttttttcc tttttttttt tttttttttt tttttttttt 7860
 tttttctcctt tttttttcct ctttttttcc ttttttttcc tttggtggct ccattcttagc 7920
 10 cctagtcaag gctagctgtg aaaggtccgt gagccgcttg actgcagaga gtgctgatac 7980
 tggcctctct gcagatcaag t 8001

15

20

25

30

35

40

45

50

55

40040	gacacagccc	gatgtggggc	gaaactctac	catagatcac	tccctctgac	ggaactactg	60
	ctcttcacga	gaaagctgac	agocattgac	ttgatatacg	tgtgtctgac	ctctcaggac	120
	ccccctccc	gagagagcca	tagtgtgtct	gacacacggt	gagtacacag	gaattgcaga	180
	gagacagcag	gagagagcca	gataccacag	ctcaatcgct	ggagatctgc	gagtcgcccc	240
	gagagagcag	taagagatga	aggtctctga	gacgtctgac	ctatgacatc	gaattgcaga	300
	gtgtgtgcga	gtgcccggtg	aaacacacag	ggcgctccat	ctgtgaacaa	gatgtgattg	360
	ctcaagaaga	aacaaacagt	ttgggtggag	ggctattcgc	ctatgacatc	gcacaacaga	420
	acagagaggt	tcgcggcgct	cgcgctgtgc	ggctgtcagc	gagagggcgc	ccggtgtctt	480
	caatgcgctg	ctctgtgact	gtgtgcctga	atgaactcga	gagagagcca	cgcgcgctat	540
	ttcttcagac	gcagactcgc	gtctctctgc	cagctgtctg	ccagctgtctg	actgaaagac	600
	ctgtagtgcg	gcacagaggg	gtctctctgc	ccagctgtctg	ccagctgtctg	actgaaagac	660
	ctgtgtgtgc	gagagagcca	gtctctctgc	ccagctgtctg	ccagctgtctg	actgaaagac	720
	ctgtgtgtgc	gagagagcca	gtctctctgc	ccagctgtctg	ccagctgtctg	actgaaagac	780
	ctgtgtgtgc	gagagagcca	gtctctctgc	ccagctgtctg	ccagctgtctg	actgaaagac	840
	ctgtgtgtgc	gagagagcca	gtctctctgc	ccagctgtctg	ccagctgtctg	actgaaagac	900
	ctgtgtgtgc	gagagagcca	gtctctctgc	ccagctgtctg	ccagctgtctg	actgaaagac	960
	ctgtgtgtgc	gagagagcca	gtctctctgc	ccagctgtctg	ccagctgtctg	actgaaagac	1020
	ctgtgtgtgc	gagagagcca	gtctctctgc	ccagctgtctg	ccagctgtctg	actgaaagac	1080
	ctgtgtgtgc	gagagagcca	gtctctctgc	ccagctgtctg	ccagctgtctg	actgaaagac	1140
	ctgtgtgtgc	gagagagcca	gtctctctgc	ccagctgtctg	ccagctgtctg	actgaaagac	1200
	ctgtgtgtgc	gagagagcca	gtctctctgc	ccagctgtctg	ccagctgtctg	actgaaagac	1260
	ctgtgtgtgc	gagagagcca	gtctctctgc	ccagctgtctg	ccagctgtctg	actgaaagac	1320
	ctgtgtgtgc	gagagagcca	gtctctctgc	ccagctgtctg	ccagctgtctg	actgaaagac	1380
	ctgtgtgtgc	gagagagcca	gtctctctgc	ccagctgtctg	ccagctgtctg	actgaaagac	1440
	ctgtgtgtgc	gagagagcca	gtctctctgc	ccagctgtctg	ccagctgtctg	actgaaagac	1500
	ctgtgtgtgc	gagagagcca	gtctctctgc	ccagctgtctg	ccagctgtctg	actgaaagac	1560
	ctgtgtgtgc	gagagagcca	gtctctctgc	ccagctgtctg	ccagctgtctg	actgaaagac	1620
	ctgtgtgtgc	gagagagcca	gtctctctgc	ccagctgtctg	ccagctgtctg	actgaaagac	1680
	ctgtgtgtgc	gagagagcca	gtctctctgc	ccagctgtctg	ccagctgtctg	actgaaagac	1740
	ctgtgtgtgc	gagagagcca	gtctctctgc	ccagctgtctg	ccagctgtctg	actgaaagac	1800
	ctgtgtgtgc	gagagagcca	gtctctctgc	ccagctgtctg	ccagctgtctg	actgaaagac	1860
	ctgtgtgtgc	gagagagcca	gtctctctgc	ccagctgtctg	ccagctgtctg	actgaaagac	1920
	ctgtgtgtgc	gagagagcca	gtctctctgc	ccagctgtctg	ccagctgtctg	actgaaagac	1980
	ctgtgtgtgc	gagagagcca	gtctctctgc	ccagctgtctg	ccagctgtctg	actgaaagac	2040
	ctgtgtgtgc	gagagagcca	gtctctctgc	ccagctgtctg	ccagctgtctg	actgaaagac	2100
	ctgtgtgtgc	gagagagcca	gtctctctgc	ccagctgtctg	ccagctgtctg	actgaaagac	2160
	ctgtgtgtgc	gagagagcca	gtctctctgc	ccagctgtctg	ccagctgtctg	actgaaagac	2220
	ctgtgtgtgc	gagagagcca	gtctctctgc	ccagctgtctg	ccagctgtctg	actgaaagac	2280
	ctgtgtgtgc	gagagagcca	gtctctctgc	ccagctgtctg	ccagctgtctg	actgaaagac	2340
	ctgtgtgtgc	gagagagcca	gtctctctgc	ccagctgtctg	ccagctgtctg	actgaaagac	2400
	ctgtgtgtgc	gagagagcca	gtctctctgc	ccagctgtctg	ccagctgtctg	actgaaagac	2460
	ctgtgtgtgc	gagagagcca	gtctctctgc	ccagctgtctg	ccagctgtctg	actgaaagac	2520
	ctgtgtgtgc	gagagagcca	gtctctctgc	ccagctgtctg	ccagctgtctg	actgaaagac	2580
	ctgtgtgtgc	gagagagcca	gtctctctgc	ccagctgtctg	ccagctgtctg	actgaaagac	2640
	ctgtgtgtgc	gagagagcca	gtctctctgc	ccagctgtctg	ccagctgtctg	actgaaagac	2700
	ctgtgtgtgc	gagagagcca	gtctctctgc	ccagctgtctg	ccagctgtctg	actgaaagac	2760
	ctgtgtgtgc	gagagagcca	gtctctctgc	ccagctgtctg	ccagctgtctg	actgaaagac	2820
	ctgtgtgtgc	gagagagcca	gtctctctgc	ccagctgtctg	ccagctgtctg	actgaaagac	2880
	ctgtgtgtgc	gagagagcca	gtctctctgc	ccagctgtctg	ccagctgtctg	actgaaagac	2940
	ctgtgtgtgc	gagagagcca	gtctctctgc	ccagctgtctg	ccagctgtctg	actgaaagac	3000
	ctgtgtgtgc	gagagagcca	gtctctctgc	ccagctgtctg	ccagctgtctg	actgaaagac	3060
	ctgtgtgtgc	gagagagcca	gtctctctgc	cc			

5
 10
 15
 20
 25
 30
 35
 40
 45
 50
 55

aaaaacttga cctgcccacc ggaactgttc cggaaagacc ccagggccac ttacaccagg 3600
 ttgtgttcgg ggctgtgact actaccata caggtcttgg 3650
 cactaccctc ggaactgcaa ctttaacatc ttcaaggtta ggaatgtacct ggggggagta 3720
 gacacagcgc tcgaagccgc atgcaattgg actcgaggag agcctgttaa cctggagacc 3780
 agggacagat cagagcttag ccgcctgctg ctgtctacaa ccgagtgcca ggtattgcc 3840
 tggctcttcca ccaacctacc ggctgtgtcc actggtttga tccatctoca tcagaacgtc 3900
 gtggagctac aatacctgta cgttataggg tcggcggttg tctcctttgc aatcaattgg 3960
 gagtatgtcc ttgtgtcctt ccttctctg gggagagcgc gcgtctgtgc ctgctgttgg 4020
 atagctgtgc tgatagctca agcagagcgc actgtgagga acctgtgtgt cctcaagcgt 4080
 gcatcogtgc ccgggcgcca tggcattctc tccctctctg ttgtctctgt ttgtcgtcgt 4140
 tcatcaaaag cgaagctggt cctggggcgc gcatatgcc tctacggcgt atggcgccta 4200
 ctccctgtcc tgtcggcggt acaacaacca gcatatgcc tggaccggga gatggcagca 4260
 tctgtcggag gcgcgggttt cgtaggtctg ataactctga ccttgtcaac gcactataag 4320
 ctgttctcgc cttaggtctat atgttggtta caatatatta tcaacaaggc cgaggcaca 4380
 ttgcaagtgt ggaatcccc cctcaacgtt cggggggggc gcatgcgt catctctc 4440
 aagtgcgaga tcaaccaga gctaatcttt acatcaaca aaatcttgtt ccgcatctc 4500
 ggtccaacta tgtgtctcca ggtgtgata caacaagtc cgtacttgt gcgcgcaca 4560
 gggctaatc gtgcattgat gctgtgcgag aagtttgtg ggggtcata ttgccaagat 4620
 gctctcatga agttggcgc actgaaggt aogtaoytt atgacatct caccocactg 4680
 cgggaactgg ccaacgcggg actacagagc ctctcgttgg cagttgagcc cgtgtctc 4740
 tctgatatgg agaccaaggt tatcacctgg ggggcagaca ccgcgcgttg tggggacatc 4800
 atcttgggoc tgcctgtctc ccgcgcgaag gggaggagga tacatctggg accgcagac 4860
 agacttgaa ggaagggttg gcgactcttc gcgcataa tccgcactact caacagagc 4920
 gaggcgtac ttggctgctc catcatgc ctccacagc cgggacagga cctgagag 4980
 gggaggtctc aagtggtctc caacgacaa caatcttctc tggcagctc cgtcaactgc 5040
 gtgtgttga ctgtctatca ttgtgcgcgc tcaaaagacc ttgcgcgcac aaaggccca 5100
 atcaaccaaa ttgataccaa ttgataccag gactctgtc gctggcaag ccccccggg 5160
 gcgctgtctc tgacacatg caactcggc agctcggacc ttatttgtt cagagagcat 5220
 gcgcatgtca ttcogtgcgc ccggcggggc gacagcagg ggaagctact ctccccagg 5280
 cccgtctctc actgaaggc ctctcgggc ggtccaactgc tctgccctc ggggagatg 5340
 ttgggcactc ttccggctgc cgttgcaacc cgaaggttgc cgaagcgtt ggaacttga 5400
 ccgcgcaggt ctatgaaac catatgcg tcccggtct tcaaggaca ctgtctctc 5460
 ccggcgtcat cgcagacatt ccaagtgggc catctacag cccctactgt tagcggcag 5520
 agcatcaagg tgcgcgtgc gtatgcagcc caagggtata agtgcttgt cctgaaccc 5580
 tccgtgcgcg ccaactag ttctgggggc tatatgtcta aggcacatgt tatcgacct 5640
 aacatcagaa ccgggttagg gaccataacc accggtgc ccataccta ctccactat 5700
 ggcaagtttc ttgcgaagg ttgttgcctc gggggcctc atgacata aatatgtat 5760
 gaagtccaat caactgactc gaccactatc ctgggcatc gcacagctc ggaacagcg 5820
 gagaagctgc gaggcgcact cgtgtgtc tcccaagta cgcctcggg atcggtaacc 5880
 gtgccacatc caaactaga ggaagtggtc ctgtccagca cttgagaat cccctttat 5940
 ggcaagacca tccccatgc gaccataag ggggggaagc aactcattt ctgcattcc 6000
 aagaaagaa gtgatgagct ccgcgcgaag ctatccggcc tggactcaat tctgttagca 6060
 tattaccogg gccttgatgt atcgtcata caactagag gagactcat ttgttagca 6120
 accgagctcc taatgacggg ctttaacggc gatttgactc cagtgtatga ctgcaataca 6180
 ggtcatcccc agacagtga cttcagcgt gaccagactc tcaacttga gacgacagc 6240
 gtgcacaaag accggtgtct acgtctgagc ccgcgaagca ggaactgtat gggagagat 6300
 ggtcattaca ggttttga tcaagagaa gggccctgg gcatgtga ttctcgtgt 6360
 ctgtgcaggt gctatgacgc gggctgtgct ttgtacagac tcaacccgc cagacactca 6420
 gttaggttgc gggcttacct aaacacacca gggctgcgc tctgcgaca ccatctggg 6480
 ttctgtggaga gcgtctttac aggcctcaac cacatagag ccaattctt tctccagact 6540
 aggcagcgag gagaacaatt cccctaacgt gtatcatacc aggtcactgt gtgcgcagg 6600
 gctcagctc caactcactc gtggagacaa atgtggaggt gtctcatag gctaaagct 6660
 accgtgacg ggcaacgcc cctgtgtgat aggtcggag ccgttcaaaa cagagtact 6720
 caaccaacc ccctcaacca aaacatagc gaatgcatc ccgttgactc ggaagcgtc 6780
 acgagcaact cgtgtgtgtc aggcagatc cttagagtc tggcgcgt taagctgaca 6840
 caaggcagcg ttgtctattgt gggcagatc atctgttgcg gaagccgcgc catctctcc 6900
 gacagggaa tcctttaccg ggaagtgat gagatgaag agtgccctc acactctcc 6960
 tcatctgaac agggaatgca gctgcgcga caattcaac agaaggaat cgggttgtg 7020
 caaacagcca ccaagcaagc ggaagctgct gctccogtgg tggaa cca gttggcyacc 7080
 ctogaagcct cctgtgcgaa gcatatgtgg aatttcaca gogggataca atatttagca 7140
 ggttttcca ctctgcctgg caaccocgg caactatcag tgaagttac cagactctc 7200
 atccacagc ccctcaacca aaacatagc gaatgcatc ccgttgactc ggaagcgtc 7260
 cgcgcacac ttgtctctcc caagcgtctc ctgtctttg taggcgcgc catctgaca 7320
 ccggtcttg gcagcatagg ccttggaag gtgcttgtg atattttgc aggttaca 7380
 cgagggtgtg cagggcgctc cgtgtcctt aaggtcatga cggcgagat gccctccac 7440
 gaggacctgg ttaactact cctgtctatc ctctccctg gccctagt cgtgggctc 7500

	gtgtgcgcag	cgatactcgc	tcggcaacgtg	ggcccagggg	agggggcgtg	gcagtgagat	7560
	aaccgcgtga	tagcgttcgc	ttccgcgggt	aaccacgtct	ccccacagca	ctatgtgtct	7620
5	gagagcgacg	ctgcagcaac	tgtcaactcag	atcctctcta	gtcttaccat	cactcagctg	7680
	ctgaagaggg	ttccaccagt	gatcaacagag	gactgtctca	cgctcagctc	cggctcgggg	7740
	ctaagagatg	tttgggatgg	gatatgcacg	gtgtgtgact	atttcaagac	ctggctccag	7800
	tccaaagctc	tgcgcgactt	ggcggagatc	cctctctctc	catgcacaag	tcggcacaa	7860
	ggagtctcgc	ggggcgcagc	catcatgcga	accaactgcc	catgtgaagc	acagatcacc	7920
	gacatgtgca	aaaacggttc	catgaggatc	gtggggccta	ggacctgtag	taacacgtgg	7980
	catggaaact	tcccccataa	ccgctaacc	accggccctc	gcaaccctcc	ccccggccca	8040
10	aattatttcta	gggcgcgtgt	gcgggtggct	gctgaggagt	acgtggaggt	tacgcggctg	8100
	ggggattttc	actacgtgac	gggcattgacc	actgacaaag	taaaagtccc	gtgtcaggtt	8160
	ccggcccccc	aattcttcac	agaagtggat	ggggctcggt	tgcacaggtc	cgtcccaagc	8220
	tgcacacccc	tctcaccgga	ggaggttcaca	ttctctgctg	ggctcaatca	atacctggtt	8280
	gggtcacacg	tcccatgcgg	gcccgaaccc	gacgtagcag	tgtctcattc	catgtccacc	8340
	gacccctccc	acattaccgc	ggagacggct	aaagctagac	tggccagggg	atctcccccc	8400
15	tcccttggcca	gctcatcagc	tagccagctg	ttctgcctct	ccttgaaggc	aacatgcact	8460
	accctgcatg	actccccgga	cgtgcacclc	atcgagccca	actcctctg	gcggcagagc	8520
	atgggcggga	acatcacccc	cgtggagtca	gaaaataaag	tagtaatttt	ggactcttct	8580
	gagccgctcc	aagcggagga	ggaatgagag	gaagtatccg	ttccgcggga	gatctccggg	8640
	aggttcaggga	aattccctcg	agcgatgcc	atatgggcac	cccgcgatta	caacccctca	8700
	ctgttagagt	cctggaagga	cccgacacac	gtccctccag	tggtacacgg	gtgtccattg	8760
20	ccgcttgcca	agcccccctc	gataccacct	ccacggaggg	agaggacgtg	tgtcctgtca	8820
	gaactctaccg	tgtctctctg	cttggcgag	ctcggccaca	agacctcggg	cagctccgaa	8880
	tctgtcgccg	tcgaacacgg	ccatgcctct	gacctctcct	ggacgccttc	cgacgcggct	8940
	gacgcgggat	tcagagctga	gtcttaaccc	acccttgagg	cccttgaggg	ggagccgggg	9000
	gatccgactg	tccacgcagc	gtcctggtct	accgtaaagc	agggagcgtg	tgagagcgtc	9060
	gtctctcgtg	cgatgtccta	cacatggaca	ggcgcccgag	tcaacgcatg	cgctcggcag	9120
25	gaacacaaag	tgcccctcaa	tgcaactgagc	aactctttgc	tcctgtcaca	caacttggtc	9180
	tatgtctaca	catctccgac	cgcgaacctg	cggcagaaga	aggtcacctt	tgacgagctg	9240
	caggctcctg	acgaaccacta	ccgggacgtg	ctcaaggaga	tgaaggcgaa	ggcgtccaca	9300
	gttaaaggcta	aactctctac	cgtggagga	gcctgtaagc	tgaaccccc	acattcgccc	9360
	agatctaaat	ttggctatgg	ggcaaaaggac	gtccgggaac	tatccagcaa	ggcgttaac	9420
	cacatccgct	ccgtgtggaa	ggacttgtct	gaagacactg	agacaccaat	tgacaccaac	9480
30	atcatcgcaa	aaaatgaggt	tttctgcctc	caaccagaga	agggggggcg	caagccagct	9540
	cgcttatcgt	tattcccaaga	tttgggggtt	cgtgtgtgcy	agaaaatggc	cctttacagt	9600
	gtggtctcca	ccctccctca	ggccgtgatg	ggctctctat	acggaattca	atactctcct	9660
	ggacacgggg	tcgagttcct	ggtgaatgcc	tggaaagcca	agaaatgccc	tatgggcttc	9720
	gcatatgaca	cccgtgtgtt	tgaactcaacg	gtcaactgaga	atgacatccg	tggtgaggag	9780
	tcaatctacc	aatgttgtga	cttggcccc	gaagccagac	aggccataac	gtcgtctaca	9840
35	gagcggcttt	acatcggggg	ccccctgact	aattctaaag	ggcgcaactg	cggtctacgc	9900
	cggtgcgcgc	cgagcgggtg	actgaacacc	agctcgogta	ataccctcac	atgttatttg	9960
	aaggcccgctg	cgccctgtcg	agctcggaag	ctccaggact	gcacgatgct	cgtatcgga	10020
	gacgaacctg	tctgtatctg	tgaagacgcg	gggacccaag	aggaacaggg	gagcctacgg	10080
	gccttcacgg	agcgtatgac	tagataactc	gccccctctg	gggacccggc	caaacacaga	10140
	tacgaacttg	agttgataac	atcatgctcc	tccaatgtgt	cactgcgcga	cgtatcatct	10200
40	ggcaaaagg	tgtaactatc	caccocgtgac	ccccaccacc	cgtctgcggc	ggctcggctg	10260
	gagacagcta	gacacactcc	agtcacttcc	tggctaggca	acatcatcat	gtatgcggcc	10320
	accttgtggg	caaggatgat	cctgatgact	catcttctct	ccactctctt	agctcaggaa	10380
	caacttgaaa	aagccctaga	ttgtcagatc	tacggggcct	gttactccat	tgagccactt	10440
	gacctaactc	agatcatcca	acgactccat	ggccttagcg	cattttccat	ccatagttac	10500
45	tctccagctc	agatcaatag	ggtggcttca	tgcctcagga	aaactggggg	accgcctctt	10560
	cgagtctgga	gacatcgggc	cagaaagtgc	ccgcgttagc	tactgtccca	ggggggggag	10620
	gctgccactt	gtggcaagta	cctctcaaac	tgggcagtaa	ggaccacagc	caaaactcac	10680
	ccaatccccg	ctcgctccca	gttgatttca	tccagctggt	gagtcacgtg	ttacagcggg	10740
	ggagacatat	atcacagcct	gtctcgtgcc	cgaccccgct	ggttctcatg	gtgcctactc	10800
	ctactttctg	taggggttag	catctatcta	ctccccaac	gatgaacggg	gagctaaaca	10860
	ctccaggcca	ataggccatc	ctgttttttt	cccttttttt	ttttcttttt	tttttttttt	10920
50	tttttttttt	tttttttttt	ttcttttttt	ttcttttttt	tttctttttt	tttctttttg	10980
	tggtctccatc	tlagccctag	tcaocgctag	ctgtgaagg	tcctgtgagc	gcttgactgc	11040
	agagagtgct	gatactggcc	tctctgcaga	tcaagt			11076

<210> 8
 <211> 8001
 <212> DNA
 <213> Hepatitis C Virus

<400> 8
 gccagccccc gattgggggg gacactccac catagatcac tcccctgtga ggaactactg 60
 tcttcaacgca gaaagcgtct agccatggcg ttgatgtgag ttgtgtgcag cctccaggac 120
 cccccctccc gggagagcca tagtgtctcg cygaaccggt gaggtaaccc gaattgccag 180
 gaagcacggg gctctctctg ggtcaacccg ctcaatgctc ggggatcttg ggtgcccacc 240
 gcaagactgc tagccgagta gttgtgggtg gcgaagaagg ttgtgtgact gctgtatagg 300
 gtcttttcgca gtgccccggg aggtctcgtta gacctgtcac catgagcagc aatcctaacc 360
 cctcaagaaa aaccaaacgt aaccaaacgg ggcgcgcgat gattgaacca gatggattgc 420
 acgcaggttc tcggcccgct ttgggtggaga ggtctattcg ctatgactgg gcaacaacaga 480
 caatcgctgt cctctgatgc gccgtgttcc ggtgtcagc gcagggggcg ccggttcttt 540
 ttgtcaagac cgaactgtcc ggtgccttca atgaactgca ggacagggca gcgcggtatc 600
 cgtgtgtggc cagcaagggc gttccttgcc cagctgtctc cgaagttgtc actgaagcgg 660
 gaagggactg gctgtctatg ggcgaagltgc cggggcagga tctctgttca tctcaactgt 720
 cctctgcgga gaaagtatcc atcatgtctg atgaatggcg gcgctgcgat agcgtgctac 780
 cggctacctg cccattcgac caccacggca aacatgcgat cgaagcagca cgtactcgga 840
 tggaaagccg tctttgtgat caggatgatc tggacgaaga gcatcagggg ctcgcgccag 900
 ccgaactgtt gcacaggttc aaggccgcga tgcccagcgg cgaggtatctc gtgtgacccc 960
 atggcgatcg ctgcttgccc aatatcatgg tggaaaatgg ccgctttttc gattctatcg 1020
 acrtgtggcg cgtgggtgtg gcggaaccgt atcaggacat agcttgggtc acccgtgata 1080
 ttgctgaaga gcttgccggc gaattgggctg accgcttctc cgtgctttac ggtatccggc 1140
 ctcccgatgc gcagcgatgc gctctctatc gcttcttctc cgagttcttc tgagttttaa 1200
 cagacacaaa cgttctccct ctacgcggat caattccggc cctctccctc cccccccact 1260
 aagcttaact gcggaagccc cttggaataa ggcgggtgtg cgtttgtcta tatgttattt 1320
 tccaccatat tgcgcttttt ttgcaatgtg agggccggga aaactggccc tgtcttctgt 1380
 acgagcatcc ctagggtgtc ttcccccttc gccaaaggaa tgaaggtctg ttggaatgtc 1440
 gtgaaggaag cagttctctc tgaagcttct tgaagacaaa caacgttgtgt agcagccctt 1500
 tgcagcgacg ggaacccccc accctggcac aggtgtcctc ggggccaaaa gccacgtgtt 1560
 taagatacac ctgcaaaagg ggcacaaccc cagtgcacag ttgtgagttg gatagttgtg 1620
 gaagagtcac aatggctctc ctcaacgcta ttccacaagg ggttcagga tgcccaagag 1680
 gtaccccatc gttatggatc tgaattgggg cctcgttgca catgctttac atgtgttag 1740
 tccaggttaa anaacgtcta gcccoccgga accacgggga cgtgttttc ctttgaaaaa 1800
 cagcataata ccatggcgcc tattacggcc tactcccaac agacgcgagg cctacttgcc 1860
 tgcatactca ctagctcac aggcccggac aggaacagg tgaaggagg gtcccaagt 1920
 gtctccaccg caacacaatc ttctctggcg accgtgttca atggcgtgtg ttggactgtc 1980
 tatcatgttg ccggtctaaa gaacctgtgc ggcocaaagg gcccaatcac ccaaatgtac 2040
 accaatgttg accaggacct cgtcgtgtcg caagcgcgcc ccggggcgcg ttccttgaca 2100
 ccaatgacct gcggcgagcc ggaaccttac ttgtcacga ggcattgcga tgcattccgc 2160
 gtgcgcggcg ggtgcgagca cagggggagc ctactctccc ccaaggccgt ctctactgt 2220
 aaggctctct cggcggttcc actgctctgc cctcggggcg accgtgtgag catctttggc 2280
 gctcgcgtgt gcaaccggag ggttgcgaag gcggtggact ttgtaccctc cagttctatg 2340
 gaaacactta tgcggtcccc ggtcttcaag gacaactcgt cccctccggc cgtaccgcag 2400
 acattccagg tgcccactct acacgcctct actggttagc gcaagacac taaggtgccc 2460
 gctcgttatg cagcccaagg gtataaggtg cttgtcctga accctgcgtc cgcgcgccac 2520
 ctagggttgc ggggtatat gtctaaagca catggtatcg accctaactc cagaactcgg 2580
 gtaagacaca tcaaccgggg tgcoccaccc actacttcca cctatggcca gttcttgcg 2640
 gacgtgtggt gctctggggg cgtctatgac atcataatat gctgatgtg caactcacg 2700
 gactgcacca ctactctgg cactcggaac gtcttgagcc aagcggagac gcttgagcgt 2760
 cgactcgtcg tgctcgcaca cgtcaacgct ccgggatcgt tcaacgtgac caatccaaac 2820
 atcagaggag ttgctctgtc cagcatctga gaaatccctc ttatggcaa agocactccc 2880
 atcagacaca tcaagggggg gaggcaacct attttctgcc attccaagaa gaaattgtat 2940
 gactctcgcg cgaagctgtc ccgctccgga ctcaatgtcg tagcatatta ccggggcctc 3000
 gatgtatcgc tcaataccac agcgggagac gtcatgtctg tagcaacgga cgtctaatg 3060
 accggcttta ccggtgaact cagctcaagt atcgactgca atacattgtt caaccagaca 3120
 gtccacttca gctctggggg cgtctatgac attgagacca cgaactgttc acaagacgt 3180
 gctcacgctc cgcagcgcg agcagacac gtgaagggca gactgggata ttacagttt 3240
 gtgactccag gagaacggcc ctccggcatg ttcatctctc cgtttctgag cagatgctat 3300
 gacgcgggct gtgcttggtt cgaactcagc cccgcgcgga cctcagttag gttcgcggct 3360
 taactaataca ccaagggtt gccctctctc caggacacac ttgagttctg gtagagcgtc 3420
 tttaacggcc tcaaccacat agacgcccat ttctgtccc agactaaaga ggcagagac 3480
 aacttccctc aactggtagc atacaggct accgtgtgag ccagggtcga ggtccacct 3540

	ccatcgtggg	accaaattgt	gaagtgtctc	atacggctaa	agcctacgct	gcacggggca	3600
	acgcccctgc	tgtagtagct	gggagccgtt	caaaaagagg	tactaccac	acacccata	3650
	accaaataca	tcatggcatg	catgtcggct	gacctggagg	tgctcacag	cacctgggtg	3720
5	ctgttaggcg	gagtcctcag	agctctggcc	gcgtattgac	tgacaacag	acagcttgct	3780
	attgtgggca	ggatcatctt	ctccggaaag	ccggccatca	ttcccgacag	ggaagtcctt	3840
	taccggagtg	tcgtagagat	ggaagatgac	gcttcacacg	ctcccttacc	gcaacaggga	3900
	atgcaagctg	cogaaacaat	caaaacagag	gcaatcgggt	tgctgcacaa	agccacccaa	3960
	caagcggagg	ctgctgtctc	cgtaggggaa	ccaaaglgac	ggacatccga	agcctctctg	4020
	gcgaagcata	tgtggaattt	catcagtgct	atacaatatt	taqacaggct	gtccacatcg	4080
	ctcggcaacc	cogcgatagc	atccatgatg	gcattcacag	ctctctacac	cagcccgctc	4140
	accaaccaac	ataccctcct	gttttaacatc	ctggggggat	gggtggcgcc	ccaacttgct	4200
	ctccccagcg	ctgctctctc	tttcgtagcg	cgccggcatg	ctggagcggc	tgttggcagc	4260
10	atagagcctg	ggaaggtgct	tgtggatatt	ttggcaggtt	atggagcagg	ggtggcagcg	4320
	gcgctcgtgg	ctcttaaggt	catgacggcg	gagatgccct	ccacocgagg	actggttaac	4380
	ctcatccctg	ctatccctctc	ccctggcgcc	ctagctcgctg	gggtcgctgtg	ccgacggata	4440
	ctgctgtcgc	acgtggggcc	aggggagggg	gctatggcgt	ggaagacgg	gctgtatgag	4500
	tcgctgtcgc	gggttaacca	cgctctcccc	acgcatcatg	tgctgtacag	cgaogctgca	4560
15	gcacgtgtca	ctcagatcct	ctctagtctt	accatcactc	agctgcagaa	gaggtctcac	4620
	catgtgatca	acgagagact	ctccacgcga	tgctcogcgt	cgtggctaa	agatgtttgg	4680
	gattggatat	gcacggtggt	gactgatttc	aagacctggc	tccagtcoc	gctcctggcc	4740
	cgattggcgg	gagtcctcct	cttctcatgt	caacgtgggt	acaaggaggt	ctggcggggc	4800
	gcagccatca	tgcaaaaccac	ctgcccattg	ggggcacaga	tcaacgggca	tgtgaaaaac	4860
	ggltctcaaa	ggatcgtggg	gcctgagcag	tgtagttaac	gctggctagg	aaactctccc	4920
20	attcaacgct	acaccacgcg	ccctctggcg	ccctcccgcg	gcacaaattc	ttctagggag	4980
	cttgtggcgg	tgctgtgtca	ggagtaactg	gaggttgatc	gggtggggga	tttccactac	5040
	tgtagcgggc	tgacaactga	cgacgtaaag	tgcccgtgtc	aggttccggc	ccccgaattc	5100
	ttcacagaag	tgagatgggt	gcggttgca	aggttaogctc	cagcgtgcaa	accctctcta	5160
	cgggaggaag	tcacatctct	ggtccggctc	aatacaatcc	tggttggttc	acagctccca	5220
	tgagagcccg	aaccggagtg	agcagtgctc	acttccatgc	tcaocgaccc	ctcccaactt	5280
25	acggcggaag	cggctaaagc	taggtctggc	aggggatact	ctcccccttt	ggccagctca	5340
	tcaagtagcg	agctgtctgc	gcttctcttg	aaggcaaatc	gcactacccg	tcatgtactc	5400
	cggcagcgtc	acctcatcga	ggccaaacct	ctgtggcgcg	agggataggg	cgggaacatc	5460
	accccgctgg	agtcagaaaa	taaggttagta	attttagact	ctttcgagcc	gctccaaagc	5520
	gaggaggatg	agagggaagt	atccgtttcc	gcggagatcc	tgccggaggtc	caggaaattc	5580
	ctccagagca	tgcccatatg	ggcagccccc	gattacaacc	ctccactggt	agagctctgg	5640
	aaggaccocg	actacgtccc	tccagtggt	caacgggtgc	caattgcgcc	tgccaaagcc	5700
30	ctcccgatac	caacttccag	gaggaaagag	acggttgctc	tgctcagaatc	taccgtgtct	5760
	tctgctctgg	cggagctcgc	caacagagac	ttcgcgagct	cugaaatcgtc	ggccgttgca	5820
	agggccacgc	caaagcgctc	tcttgaccac	ccctccagcg	acggcagcag	gggataccag	5880
	gttgatcgct	actcctccat	gcctcccctt	gagggggagc	cgggggagtc	cgatctcag	5940
	gaocggctct	ggctacagct	aagcagagag	gctagtggag	acgtgtctgt	ctgctcgatg	6000
	tctccacatc	ggacagcgcc	ctgtgatcag	ccatgcgctg	cggaggaacc	caagctgcgc	6060
	atcaatgcac	tgagcaactc	tttgcctcgt	caccacaact	tggtctatgc	tacaaactct	6120
35	cgcagcgcac	gctcgcggga	gaagaaggtc	acctttgaca	gactgcaggt	cctggagcac	6180
	catctacogg	acgtgtcna	ggagatgaag	gcgaaggcgt	ccacagttaa	ggctaacttt	6240
	ctatccgtgg	aggaagcgtg	taagctgcag	cccaccactt	cggcagagct	taatttggc	6300
	tatggggcca	aggacgtccg	gaacctatcc	agcaaggccg	ttaaccacat	cogctccgtg	6360
	tggaaggaat	tgctgggaag	cactgagaca	coaattgaca	ccaactcat	ggcaaaaaat	6420
	gaggttttct	gcttccaaac	agagaagggg	ggccogaagc	cagctcgctc	tatcgtatcc	6480
40	ccagatttgg	gggttgtctgt	gtgcagaaaa	atggcccttt	acgatgtggt	ctccacccctc	6540
	ctccagccgc	tgatgggtctc	ttcatacgg	ttccaaactc	ctcctggaca	cggggtcgag	6600
	ttccttggtg	atgctcgaaa	agcgaagaaa	tgccctatgg	gcttcgcata	tgacaaccgc	6660
	tgctttgact	caaacggtcac	tgagaaagac	atccgtgtgt	agggagcaat	ctgtaaatgt	6720
	tgtagctgct	ccccgaagac	cagacagccc	ctccagagct	gctttacatg	gctttacatg	6780
	ggggggcccc	tgactaatc	taaaaggcag	aactcggctc	atcgccgtgt	cgcgcgcagc	6840
	gggtgactga	cgacacagct	cgttaataac	ctcacatggt	acttgaagac	cgcctgogcc	6900
45	tgctcagctg	cgaagctcca	ggactgcagc	atgctcgtat	gcggagagaca	cttctgtttt	6960
	attctgtgaa	gcggggggac	ccaagaggac	gagggcgagc	taogggcctt	cacggaggct	7020
	atgactagat	actctgcccc	ccctggggac	cgcgccaaac	cagaataaga	cttggagttg	7080
	ataacatcat	gctctcccaa	tgtnctcagt	ggcagcagat	catctggaca	aaaggtgtac	7140
	tatctacccc	gtgacccctt	gcccggggtg	gcgagggtgt	cgtggagagc	agctgagaac	7200
	atcccatgca	attctcggct	agcaaacatc	ataatgtatg	cgcacacctt	gtggcagagg	7260
	atgatcctga	tgactcatct	cttctccatc	cttctagctc	aggaacaact	tgaaaaagcc	7320
50	ctagattgtc	agatctacgg	ggcctgttac	tccattgagc	cacttgacat	acctcagatc	7380
	atttcaacgc	tccatggcct	tacgcatttt	tcactccata	gttactctcc	aggtgagatc	7440
	aatagggtgg	cttcatgcct	caggaaacct	ggggtaccgc	ccttgcgagt	ctgagagatc	7500

EP 1 043 399 A2

5 cgggccagaa gtgtccgccc taggctactg tcccaggggg ggagggctgc cacttggtgc 7560
 aagtacctct tcaactgggc agtaaggacc aagctcaaac tcaactccat cccggctgcg 7620
 tcccagttgg atttatccag ctggttcggt gctggttaca gcgggggaga catatatcac 7680
 agcctgtctc gtgcccgacc cgcctggttc atgtggtgcc tactcctact ttctgtaggg 7740
 gtaggcacat atctactccc caaccgatga acggggagct aaacactcca ggccaatagg 7800
 ccatacctgt tttttccctt tttttttttc tttttttttt tttttttttt tttttttttt 7860
 ttttctcctt ttttttctt ctttttttcc ttttctttcc tttggtggct ccattcttagc 7920
 10 cctagtcacg gctagctgtg aaaggtccgt gacccgcttg actgcagaga gtgctgatac 7980
 tggcctctct gcagatcaag t 8001

15

20

25

30

35

40

45

50

55

<210> 9
 <211> 11076
 <212> DNA
 <213> Hepatitis C Virus

5

<400> 9
 gccagcggccg gattgggggc gacactccac catagatcac tcccctgtga ggaaactactg 60
 tcttcaagca gaaagcgctct agccatggcg tagtatgag tctgctgcaag cctccaggagc 120
 ccccccctccc gggagagcca tagctggctcg cggaaaccggt gaggtaacacg gaattgcacag 180
 gaagcagcggg tcttttcttg gatacaacgg ctcaatggcct ggagatttgg gcgtgcggccc 240
 10 gcgcgactcg tagcgcagta gtgttggttc gcgaaaggcg ttgtgttact gcctgatagg 300
 gtgccttcgca gtgcggccgg aggttctcgt gacgcgtgac catgagcagc aatcctaatac 360
 ctcaagaaaa aacccaacyt aacaccaag gcgcggccat gattgaacaa gatggattgc 420
 acycaaggttc tccgcgcgct tgggttgaga ggctatttcg ctatgactgg gcacaacaga 480
 caatcggtcg ctctgatgcc gccgtgttcc ggctgtcagc gcaggggcgc ccggttcttt 540
 15 ttgtcaagac cgaactgtcc ggtgcccctga atgaactgca ggaaggagca gcgcgctat 600
 ogtgtctagc cagcagcggc gtctcttagc cagctgtgct cgaactgttc actgaaagcg 660
 gaaggagctg gctgtattg ggcgaagtg gcggggagga ttctctgca totcaacttg 720
 ctcttcgcga gaagatattc atcatggctg atgcaatgag gcgcgtgat acgcttgatc 780
 cgtctactcg ccattcgac caccaaagca aacatcgcat cgagcagca cgtactcgga 840
 ttgaaagcggg tcttctcgat caggatgata tggaggaaga gcatcagggg ctgcgcgca 900
 cgaactgttt cgcacggctc aaggcgcga tgcccagcg caggagatct ctgtgaccc 960
 atggcgatcg ctgcttgccg aatatcatgg tggaaaaatg ccgcttttct ggattcatcg 1020
 20 actgtggccg gctgggtgtg gggagccgct atcaggacat acgcttggct accgtgata 1080
 ttgctgaaga gctggcggcg gaatgggctg acgcgttctc cgtgctttac ggtatcgcg 1140
 ctcccgatgc gaagcgatc gctctctatc gccttcttga agacttttaa tgaattttaa 1200
 cagaccacaa cggtttctct cttagcggat caattccgcg cctctccctc cccccccct 1260
 aacgttactg gccgaagcgg cttggaataa ggccggtgtg cgttttgtata tatgttattt 1320
 tccacacata tgcgcttctt tggcaatgtg agggcccga aacctggccc tgtctctctg 1380
 25 acgagcattc cttaggggtct tctccctctc gccaaaggaa tgaacaggtt gttgaatgtc 1440
 ttgaaggaa cagttctctct ggaagcttct tgaagacaa caacgtctct agcagacctt 1500
 ttgcaggcagc ggaacccccc acctgcggac aggtgctctc gcggccaaaa gccacgtgta 1560
 taagtatac ctgcaaaagc ggacaaaccc cagtgccacg ttgtgagttg gatagttgtg 1620
 gaaaagatca aatgtctct ctcaaacgtta ttcaaaagc ttgtgaaagg tgcacagaa 1680
 tgaccacatt gtatggagat tgatctgggg cctcgttgca catgctttac atgtgttaa 1740
 30 tgcagattaa aaaaagctcta ggcggccgca accacgggga cgtgttttct ctttgaaaaa 1800
 cagcataata ccatgggcac gaatctaaa cctcaagaa aaaccaaacg taacaccaac 1860
 cgcgcgccac aggaactcaa gtctccggcg ggtgttca gttcgtgtg agtttaacct 1920
 ttgcgcgcaa gggggcccaa gttgggtgtg gcgcgacta gaaagacttc cagacggtag 1980
 caactcctgt gaagcgacaa acctatcccc aagctctgac acccgagggg tagggcctg 2040
 35 gctcagcccg ggtaccctcg gccctctatc ggcaatgagg gcttgagggt gccagagtag 2100
 ctctctgca cccgtggctc ttggctagt ttggggcccg ggcagccccc gctgagtag 2160
 cgcacttttg gtaaggtcat cgtataccct acgtggcgt tgcgcgatc calggggtac 2220
 attcgcctcg tggcgccccc cctaggggcg gctgcccagg ccttgccgca tggcgtccg 2280
 gtctctgagg acggcgtgaa ctatgcacaa ggaactctgc ccggttgctc ctctttctac 2340
 ttcttttttg ctttctgttc ctgtttgacc atccagactt ccgcttatga agtgcgcac 2400
 gtatccggag tttactcagt cagcaacgac tgcctcaacg caagcattgt gtatgagca 2460
 40 gcggacatga tcatgcatac ccccggtgct gtgcctcgct ttcgggagaa caactctcc 2520
 cagctctggg tagctgtcac tcccacgctc gggccagca acgtagcgt ccccactcgt 2580
 acgatacag gccatgtcga ttctgtctc gggggccgct cctctgtctc gcctatgac 2640
 ttggagatgc ttctgggcat tgttttctc gtgcgcagc tggtaacctt ctgcgctcgt 2700
 cggccagaga cagtacaga ctgcaattgc tcaatata ctccgcacgt gacacgtcac 2760
 cgtatggctt gggatatgat gatgaactgg tcaactacag cagccctagt ggtatcgac 2820
 ttaactcgga tcccaacagc tgtctggat atgttgagg ggcccatgt gggagtcta 2880
 45 gcggcgcttg cctactattc catggtgggg aactgggcta aggttctgat ttgtatgta 2940
 ctctttcgcg gogttgcagg gggaaacct atgacaggg gaaagatggc caaaaaccc 3000
 ctcgagatta cgtccctctt ttaccccggg tcatccaga aaatcagct tgtataaccc 3060
 aacggcagct ggacataca caggactggc ctgaaactgca atgactctct caacactggg 3120
 50 tctctgtg cgtgttcta cgtcaacgca ttcaactcat ctggaatgccc agagagcag 3180
 gccagctgca gccccataga cgcgttgcct cagggtgtgg gccccatcac ttacaaatg 3240
 tcaacagct cggaccagag gcttattgt ttgcactag caccocggcg gtgcggtatc 3300
 ttaaccggcg cgcaggtgtg ttgtccagtg tactgtctca ccccaagccc tgtcgtgtg 3360
 ggaagcagcg accggttcgg cgttccatag tacaagtgg ggaagaaatg gacggcagct 3420
 ctgctcttta caacaacgcg ccgcgcgcaa gcaactggt ttgctgtatc atggaatgat 3480
 agcatgggtt tccaagaac gtgcggggcg ccccgtgta aactcggggg gatcgcgcaat 3540

55

aaaaacttga cctgcgccac ggcactgttc cggagcacc cagaggcac ttacaccaag 3600
 tggctgttgg ggccttgggt gacacocaga tgcctgttgc atccacata caggctttgg 3660
 cactaacocct gcactgtcoa ctttaacatc ttcaagytta ggcattagct ggggggagtg 3720
 gagcaacagg ctcagagcgc atgcacattg actcagggag agcgtttgtt cctgagagac 3780
 agggacagat cagagcttag ccocgtgctg ctgtctacaa cggattggca ggtattgcc 3840
 ttgtctctca coaacctaac ggcctgtctc actggtttga tccatctoca tcagaaogtc 3900
 gtggacgtac aataccctga cgttataggg tcggcggttg tctctcttgc aatcaaatgg 3960
 gaggatgtcc tgttgcctct ccttctctgc gggagcggcg gcgctgttgc ctgcttggtg 4020
 atgatgtctc tgaatgtcoa agctgaggcg ggcctagaga acctggtgtg cctcaacgtg 4080
 gcatccgtgc ccggggcgca tggcattctc tctctctcgt tgttctctgt tgcgtccgtg 4140
 tacatcaagg gcaagctggg cctcgggggc gcatatgcc tctacggcgt atggccgcta 4200
 ctctcgtctc tgcgtggcgt accacaogca gcatagccca tggacccggg gatgcagca 4260
 tctgtggagg gcgcggtttt cgttagcttg atactcttga ccttgtcaac gcactataag 4320
 ctgttctcgt ctaggctcat atggttggtta caatatitaa tcaocaggcg cagggcacac 4380
 ttgcaagtggt ggtaccoccc cctcaacgtt cggggggggc gcgctgcgt catctctctc 4440
 acgtygcgga tcaacocaga gctaatcttt accatcacca aaacttgtgt cgcacactc 4500
 ggtccactca tgggtgtcoa ggcgtgata accaaagtgc cgtacttgtg gcgcgcacac 4560
 gggctcattc gtgcatacat gctggtgcgg aaggttgtgt ggggtcatta tgtcaaatg 4620
 gctctcatga agttggcgcg actgacaggt actgactytt atgacacatc caccocactg 4680
 gctactctgg ccaacggcgg cctacagac ctgcggtggc cagttgaagc cgtcgtcttc 4740
 tctgatatgg agaccaaagt tatcaactgt gggcgagaca ccgcgcgtgt tggggacatc 4800
 atcttggggc tgcocgtctc cgcocggagg gggagggaga tacatctggg acocgcagac 4860
 agccttgaag ggcgggggtg cagcactctc gcgcctatta cgcctactc gacacagag 4920
 cggagctatc ttgctctatc catcactagc ctcaacggcg gggacaggaa ccaagctgtg 4980
 gggagagctc aagtgtctc ccacgcacaa caattcttcc tggcgccctg cgtcaatgac 5040
 gtgtgttggg ctgtctatca tggctccggc tcaaaagccc ttgcgcoccc aaaggggcca 5100
 atcaacccaaa tgtacaocaa tgtgacaag gacctoytgc gctgcgaagc gcccccgggg 5160
 gcgcgttctc tgaacacatg caactcggcg agctcggacc ttacttgtgt cacagagcat 5220
 gcgatgttca ttcocgttgc ccggcgggcg gacagcaggg ggcgactact ctcccccagg 5280
 ccgcctctct actgaaggg cctctcggcg ggtccactgc tctgcocctc ggggcaagct 5340
 gtgggcatct ttgcggtcgc cgttgcaccc cggaggggtg cggaggggtg ggaacttga 5400
 cccttcagat ctatggaaac cactatgcgg tccccgtctc tcaogggaca ctgcctccct 5460
 ccgcgcgtac cgaacacatt ccaggtggcc catctacag ccctactgtg tagcggdaag 5520
 agcaactaag tgcocgtcgc gtatgcagcc caagggtata agtgctgtgt cctgaacocg 5580
 tccgtgcgcc ccaacctagg tttcggggcg tatatgtcta aggcacatgt tatcgacct 5640
 aacatcagaa tcggggtaag gacacatccc acgggtgccc ccatcaagta ctccacctat 5700
 ggcacgtttc tgcgaagcg tgggtgctct gggggcgccct atgacatcat aatatgtgat 5760
 gagtgcactc caactgactc gacactatcc clgggcatcg gacagctccl gacacagcg 5820
 gagaagcgtg gaggcgagct cgtcgtctc cgcacaccta cgcctccggg atcgggtcac 5880
 gtgccacatc caaacatoga ggaagtggtct ctgtccagca cttggagaat ccccttttat 5940
 ggcaaaagcca tccocatoga gacacatcag gggggaggcg acctoatttt ctgcactcc 6000
 aagaagaat gtgatgagct cgcgcogaag ctgtccggcc tcggactcaa tgcgttagca 6060
 tattaccggg gcottgatgt atccgtcata coaactagcg gagagctcat tgcgttagca 6120
 acggagctct taatgacggg cttaccggtc gacttcgact cagtgatoga ctgcaataca 6180
 tgtgtcaccc agacagtoga cttacagctc gaccogactc tcaacatga gaagagacac 6240
 gtgtccaag cgcggtgtc acgtccagcg agcagaggca gaaclgttag gggcaggtg 6300
 gcatitlaca ggttttga tccagagaa cggccctcgg gcatgttoga ttctcgtgt 6360
 ctltgcgagt gctatgacg ggcctgtgct tgtatcagc tcaagccgc cagacactca 6420
 gttaggttgc ggccttacct aaaaacacca ggggtccccc tctgcagga ccaatctgag 6480
 tctcgtggaga gcgtctttac aggccttaac cacatagac ccaatttctt gtccagact 6540
 aagcaggcag gagacaactt cccctacctg gtatgcatac agactacgtg gtgcgcaggg 6600
 gctcaggctc caactccatc gtgggaccaa atgtggaagt gctctatacg gctaaagct 6660
 agcgtgcagc ggcacaagcc cctgctgat aggcctggag ccgttcaaaa cgaagttact 6720
 accacacccc ccaataccaa ataacatccc gctatgttga cgcgtcactc gaggctgtg 6780
 acgagcactc ggtgcttgt aggcgagtc cttagcgtct tggcccgcta tttcctgaca 6840
 acagcgagcg tggctatitl gggcagagtc atcttctcgc gaaagccgcg catcatccc 6900
 gacagggaag tcttctaacc ggaattgcat gagatggaag agtcgcctc acacctccc 6960
 tatcatogaac aggggaatga gctgcogaaa caattcaaac agaaggaat cgggttctgt 7020
 caaacagoca coaagcaagc ggaagctgct gctccogtgg tggaaatcaa gtggcgacc 7080
 atcgaagcct tctggggcaa gcatatgtgg aatttcaata cggggataca atatttagca 7140
 gctgttgcga ctcgcgcgg ctcgcgcg atagactcac tgatggcatt caoagcctt 7200
 atcaacagcc cgttcacac ataacatccc gctatgtta acactctgg gggatgggt 7260
 ccgcoccaac ttgctctctc cagcgtctgt tetgttttgc taggcgcggc calcgtgga 7320
 gcgcgtgttg gcagcatagg ccttgggaag gtcttgttg atattttgc aggttatgga 7380
 gcagggggtg caggcgcgct cgtgcctttt aaggtcatga cggcgagat gccctccac 7440
 gaggacctgg ttaacctact cctcgtctat cctctccctg gcgcctagct cgtcggggtc 7500

	gtgtgcgcag	cgatactcgc	tgcgcacgtg	ggcccagggg	agggggcgctg	gcagtggaat	7560
	aaccggctga	tagcgttcgc	tgcgcgggtg	aaccacgtct	ccccacacga	ctatgtgcct	7620
	gagagcgacg	ctgcagcagc	tgtcactcag	atcctctcta	gtcttaccat	cactcagctg	7680
5	ctgaagaggc	tccaccagtg	gatcaacag	gactgtccca	cgcctatgctc	cggctcgtgg	7740
	ctcaaggatg	tttgggattg	gatatgcag	gtgtgactgc	atttcaagac	ctggctccag	7800
	tccaaagctc	tgcgcgattc	gcccggagtc	cctctttctc	catgtccaag	tgggtacaag	7860
	ggagtcgtgc	ggggcgacgy	catcatgcac	accaoctgccc	catgtggggc	acagatcacc	7920
	ggacatgtgc	aaaaacggtc	catgaggaatc	gtggggcctc	ggacotgtag	taacacgttg	7980
	catggaacat	tcccacttaa	cgcgtacacc	acgggcccct	gcaagccctc	cccgccgcca	8040
	aattatttcta	gggcgctgtg	ggcggtggct	gctgaggagt	acgtggaggt	tacgcgctgt	8100
	ggggattttcc	actacgtgac	ggcctagacc	actgacgaag	taaaagtccc	gtgtcaggct	8160
	ccggcccccc	aatttctcac	agaagtggtg	ggggtgoggt	tgccagagta	cgctccagcg	8220
	tgaacacccc	tctcaecgga	ggaggtcaca	ttcctgggtcg	ggctcaatca	atacctggtt	8280
	gggtcacagc	tcccattgca	gcgcgaacog	gatgtagca	tgcctacctc	catgtccacc	8340
	gacccctccc	acattacggc	ggagacggct	aagcgtaggg	tggccacggg	atctctctcc	8400
15	cccttggcca	gctcatcagc	tagcacgctg	tctgcgcttc	ccttgaaggc	aacatgcact	8460
	accctgcatg	actccccgga	cgtgaccttc	atcgaggcca	acctoctgtg	cgggcagagc	8520
	atgggcggga	acatcacccg	cgtggagtca	gaaaataagg	tagtaatttt	ggactctttc	8580
	gagcgctccc	aagcgggaga	ggatgagagg	gaagtatccc	ttcccgcgga	gatcctcgcc	8640
	aggtccaggga	aattcctcgc	agcagtgccc	atatgggcac	gcccggatta	caacccctca	8700
	ctgttgagat	cctgggaagg	ccccgactac	gtcctccag	tggtaacgg	gtgtccattg	8760
20	ccgccttgcca	agggccctcc	gataccacct	tcaaggagga	agaggacgtt	tgtccttgca	8820
	gaactctacgc	tgtcttctgc	cttggcgagg	ctcgccagca	agacottggc	cgactcgcga	8880
	ttctcggggc	tcgcagacgy	caagcgcaacg	cgtctctcgt	accagccctc	cgagcagcag	8940
	gacgcgagtc	ctgcactcca	ctctgacccc	cccttgagga	ggaacggggc	ggagccgggc	9000
	gatcccgatc	tcagcgacgy	gtcttggtct	accgtaaagc	agggagctag	tgaagcgctc	9060
	gtctgctgct	cgatgtccta	caatcggaca	ggcgccctga	tcacgcctag	cgctcgagag	9120
25	gaaaacaaagc	tgccctctca	tgcactgagc	aactctttgc	tcctgcacca	caactgtgtc	9180
	tatgtctaca	catctcgcag	cgcaagcctg	cggcagaaga	aggtcacctt	tgcacgactg	9240
	caggctcctg	acgacccaat	cgggacgtg	ctcaaggaga	tgaagcgcaa	ggcgctacca	9300
	gttaaaggcta	aacttctatc	cgtggaggaa	gcctgtaaag	tgaaccccc	acatttccgc	9360
	agatctaaat	ttggctatgg	ggcaaaagac	gtccggaaac	tatccagcaa	ggccgttaac	9420
	caatcccgct	cgtgtggaaa	ggacttgctg	gaagacactg	agacaccaat	tgacaccaac	9480
30	atcatgagcaa	aaaatgaggt	tttctcgtgc	caacacagaa	agggggggcg	caagccagct	9540
	cgcccttatcg	tattccccaga	tttgggggtt	cgtgtgtgcy	agaaaatggc	cctttacgat	9600
	gtgtgtctcca	ccctccctca	ggcgtgatg	ggctcttcat	acggaattcca	atactctcct	9660
	ggacagcggg	tcaggttccct	ggtgaatgcc	tggaaagcga	agaaatgccc	tatggcgctc	9720
	gcatatgaca	cccgtctgtt	tgaactcaag	gtcaactgaga	atgacatcag	tgttgagaga	9780
	tcaatctaac	aattgttgtga	cttggccccc	gaagccagac	agggccataag	gtcgctcaac	9840
35	gagcgctgtt	acatcggggg	ccccctgact	aattctaaag	ggcgaagctg	cggtcatacg	9900
	cgtgtcccggt	cgagcggtgt	actgacgacc	agctcggtta	atacctcac	atgttacttg	9960
	aaggccgctg	cggcctgtgc	agctgcgaag	ctccaggact	gaacgatgct	cgtatcgga	10020
	gacgaacctg	tcttattctg	tgaagacgcg	gggacccaag	aggaagaggg	gagctctacg	10080
	gacttcacgg	aggetatgac	tagatactct	gcccctcgtg	gggacccggc	caacccaagc	10140
	tacgaactgt	agttgatatc	atcatgctcc	tccaatgtgt	cagtccgcga	cgatgactct	10200
40	ggcaaaaggg	tgtactatct	caaccgtgac	cccaacccc	ccttgcgcg	ggctcgtgtg	10260
	gagacagcta	gacacacctc	agtcacattc	tggctaggca	acatcatcat	gtatgcgcc	10320
	accttgtggg	caaggatgat	cctgatgact	catttctctc	ccatctctat	agctcaggaa	10380
	caacttgaaa	aagccctaga	ttgtcagatc	tacggggcct	gttactccat	tgagccaatt	10440
	gacctacctc	agatcattca	acgactccat	ggccttagcg	catttctcat	ccatagtctac	10500
	tctccagctg	agatcaatag	ggtggcttca	tgcctcagga	aaactggggg	accgcctctt	10560
45	cgagtctgga	gacatcgggc	cagaagtgtc	cgcgctagcc	tactgtccca	ggggggggag	10620
	gctgcacatt	gtggcaagta	cctcttcaac	tggggcagtaa	tgaccaagct	caaatcactc	10680
	ccaatccccg	ctgcgtccca	tctggattta	tccagctggt	gttctgtctg	ttacagcggg	10740
	ggagacatat	atcacagcct	gtctcgtgccc	cgaccccgct	ggttctatgt	gtgcctactc	10800
	ctactttctg	taggggttag	catcatctca	ctcccccaac	gatgaacggc	gagctaaaca	10860
	ctccagggcca	ataggccatc	ctgttttttt	cccttttttt	ttttcttttt	tttttttttt	10920
50	tttttttttt	tttttttttc	ttcttttttt	ttctcttttt	tttctttttc	tttctttttg	10980
	tggctccatc	ttagccctag	tcaacgctag	ctgtgaaagg	tcctgtagcc	gcttgactgc	11040
	agagagtgct	gtactcgcc	tctctgcaga	tcaagt			11076

<210> 10
 <211> 8001
 <212> DNA
 <213> Hepatitis C Virus

5
 <400> 10
 gccagc0000 gattgggggc gacactccac catagatcac tcc00ctgga ggaactactg 60
 tcttca0cga gaaagcgctt agc0actggcg ttagtatgag tgt0gtgcag cctccaggag 120
 ccc0cc0ccc gggagagc0ca tagtgcctgc cggaa0cggg gagcacacgc gaattgcccag 180
 ga0c0ccggg tctcttcttg gatcaac0cg ctcaa1tgct gga0atttgg g0gtgc0ccc 240
 10 g0cgagcctc tagc0agatg ttgttggttc gc0aa0agcc tftgtgtgat gctgtatagg 300
 gt0ctttg0a gtgc0cc0gg aggtct0gta gac0ctgcac catgag0cag aatccta0aa 360
 ctcaaa0aaa aa0caa0cgt aacaccaa0g gg0cg0ccat gattga0caa gatgattgc 420
 acg0aggttc t00gg0ccgt tgggtggaga ggctatt0gg ctatga0ctg gcacaacaga 480
 caat0cg0gt ctctgatgcc gc0gtgttcc ggctgtcagc gca0gggc0gc c0ggttcttt 540
 ttt0ca0agc 0gac0ctgtcc ggtgc0ctga atgaactgca gga0gagcca gc0cg0ctat 600
 15 0gtgtgcagg ca0ga0gggc gtc0cttggg cagctgtgct cga0gtttgc actgaag0gg 660
 gaagga0gtg gctg0tatgg ggc0ga0gtg 0gg0gg0ga tct0ctgtca tctca0cttg 720
 ct0ctgc0ga gaaagtatcc atcatgggtg atgcaatgcg 0gc0gtgc0at a0c0ctgatc 780
 0ggtc0ac0t cccatt0gac ca0caa0cga aacat0gcat 0ga0gag0ca 0gtact0gga 840
 tggaa0ccgg tctt0tg0at caggatgatc tgg0ca0ga gcat0agggg ct0cg0ccag 900
 ccgaactgtt gc0caggctc aggc0gc0ga tg0cc0agc 0gagatctc gt0gtgacc 960
 atg0cgatgc ctgcttgc0g aatatcatgg tga0aaatgg c0gcttttct ggattc0atg 1020
 20 actgtgg0cg gctgggtgtg 0gcga0c0ct atcaggacat ag0cttggct acc0gtgata 1080
 ttgctga0ga gcttgg0ggc gaatgg0gtg ac0gctt0ct 0gtgctttac ggtat0cgg 1140
 ct0c0gattc gcag0gc0tc g0ctct0ctc ac0t0ctga 0ga0ttcttc tga0ttttaa 1200
 gac0ca0caa 0ggtt0ct0ct ct0ag0ggat caat0cc0cc 0ct0c0ctc 0cc0cc0ctc 1260
 aac0gttact gc0ga0gcc 0ttg0aata gg0c0gtgtg 0gttt0tcta tatgtatttt 1320
 25 t0ccac0at tg0cgtt0tt tggcaatgtg aggg0cc0ga aac0tgc0cc tgt0t0ttg 1380
 acgagcattc ct0gggt0ct t0cc0ct0tc gc0aa0gaa tga0aggtct tgtgaatgc 1440
 gtga0ga0ag cagtt0ct0ct gga0agttct tga0gacaaa caacgtctgt ag0cacc0ct 1500
 tg0cag0cag gga0cc0ccc 0c0tgg0gac aggtg0ctct 0gcgg0caaaa gcc0c0gtga 1560
 taagat0aac ctgcaaa0gg gc0caaa0cc 0agtgc0acg tttgagtgtg gatagtgtgtg 1620
 gaa0ag0ta aatgtgt0tc ct0aa0g0ta ttc0aa0agg ggt0ga0ga tgc0ca0ga 1680
 gt0cc0aat gtatgggatc tgatctgggg 0ct0gtg0ga catgttttac atgtgtttag 1740
 30 t0caggttaa aaaa0gt0ta gg0cc0c0ga acc0ag0gga 0gtgttttct ctttga0aaa 1800
 cc0gataata ccatgg0gc0 tattac0gcc tactcca0ac agac0cg0ga 0ctacttgc 1860
 tgcatcatca ct0agc0t0ac aggc0gg0ag agga0ccagg t0cag0ggga ggtccaagt 1920
 gtct0cc0cg ca0aca0atc ttt0ctgg0c 0c0tgc0tca atggt0gtgt ttggac0tgc 1980
 tatcatgggt 0cg0gt0caa gac0cttgc0 gg0ccaa0gg gcc0aat0ac 0caa0atgac 2040
 acca0gttgg accagga0ct 0gt0gc0tgy ca0gc0cccc 0c0gg0gg0cg ttc0ttg0ca 2100
 35 0catg0ac0t gc0gg0g0tc gga0ctttac ttgg0ca0ga ggt0tg0ga tgc0at0cg 2160
 gt0gc0c0gg 0gg0gc0a0g ca0gg0gg0gc 0ta0ct0ccc 0ca0g0c00gt 0ct0actgt 2220
 aag0gt0ctt 0gg0gg0tcc 0ctg0ct0gc 0c0t0gg0gc 0ag0ctgtgg cat0ctt0gg 2280
 gctgc0gtgt gca0cc0ag0 ggttgc0aag 0gc0tgg0act ttgtacc0gt 0ag0ct0atg 2340
 gga0c0acta tg0cgt0ccc ggtcttca0g gaca0ct0gt 0c0ct0c0gc 0gtacc0cag 2400
 acatt0c0ag tgg0cc0atc ac0gc0cc0t 0ctggt0ag0 gca0ag0cac ta0ggtg0cg 2460
 gctgc0tatg cagg0ca0ag gtataag0tg 0ttgt0ctga 0ac0gt0c0gt 0c0gc0caac 2520
 40 ct0agtttga ggg0gtatat gt0cta0g0ca catgtat0g acc0ta0cat caga0at0ggg 2580
 tga0g0ca0ca tca0cc0ggg tg0cc0c0tc 0cgt0act0ca 0ct0atg0ga gttct0ctgc 2640
 gac0gtgtgt gct0c0ggg 0gc0ctatgac atca0ta0at gcatgt00gt 0ca0t0ca0ct 2700
 gact0ca0ca 0tat0c0ggg 0at0c0g0ca gt0ctg0aac 0gc0ag0agc ggtc0g0agc 2760
 0g0act0gtc tgct0gc0ac 0g0ta0gc0ct 0c0gg0at0cg tca0c0tgc0 acat0cc0aa 2820
 att0ag0agg ttg0ct0gtc 0ag0act0ga gaa0t0cc0t tt0atg0caa ag0cat0ccc 2880
 at0ga0g0ca tca0gg0ggg gagg0ca0ctc atttt0tgc0 att0ca0gaa gaa0atgtat 2940
 45 gagct0c0cg 0ga0agctgtc 0gc0ct0cga 0tcaatgtgt tagcatatta 0cgggg0c0ct 3000
 gtatgat0cg tca0ac0ac tag0gg0agc gtcattgt0g tagca0c0ga 0cgt0cta0t 3060
 a0gg0cttta 0c0gg0ac0ct 0g0ct0c0gt at0ga0ctga atacatgtgt ca0cca0gaca 3120
 50 gt0c0ct0ca g0c0ga0ccc 0ac0tt0acc attg0ag0ga 0gac0gt0gc 0ca0ag0cg 3180
 gct0c0c0ct 0gc0ag0c0gg ag0c0ag0act ggt0ag0gga ggt0g0g0cat ttac0agttt 3240
 gt0cat0cag gaga0cgg0c 0t0cgg0catg tt0catt0ct 0cgttt0ctgt 0cagtg0tat 3300
 gac0c0gg0ct gt0ctt0gta 0ga0ct0c0g 0c0c0cg0ga 0c0catt0gt gtt0c0gg0ct 3360
 ta0cta0aaa ca0c0ag0gtt gc0c0gt0cgc 0ag0ac0atc ttgagt0ctg gaga0gc0tc 3420
 tttac0agcc tca0cca0cat ag0c0cca0ct t0t0t0ccc ag0cta0gca gga0c0g0ac 3480
 aact0cc0ct 0c0tgg0tag at0ca0ggct ac0gtgt0cg 0c0aggt0cga ggt0c0ac0ct 3540

	ccatcgtggg	acccaaatgtg	gaagtgctct	atcgggttaa	agcctacgct	gcacgggcca	3600
	acgcctccctgc	tgatatagctg	ggagagcgtt	caaaaacaggt	tactaccacc	acacccccata	3660
	acccaataca	tcattgccaat	catgtccgct	gaacctggagg	tgctccacag	caacctgggtg	3720
5	ctggtaggcg	gagtcctcagc	agctctggcc	gcgtattgcg	tgacaacagc	gacgcttgctc	3780
	atttgtggcca	ggatcatcttt	gtccggaaag	cgggcacata	ttcccgacag	ggaagtcctt	3840
	taccggagtg	tcgatgagat	ggaagatgct	gcctccaccc	tcctctacat	cgaacaggga	3900
	atcgagctcg	cogaaacaat	caaacagaa	gcaatcgggt	tgctgcgaac	agccaccaag	3960
	caagcgaggg	ctgctgtctc	cgttggtgaa	tccaagtggc	ggaacctaga	agcctctcgg	4020
	gcgaagcata	tgtgaaatt	catcacagg	atacaattat	tagcagagct	gtccacata	4080
	ccctgcacac	cogcgatagc	atcaactgat	gcattccacg	ctctctatac	cagcccgctc	4140
10	acccaacca	ataccctcct	gttttaacatc	ctggggggat	gggttgccgc	ccaacttgct	4200
	ctcccccagcg	ctgctctctgc	tttcgtagcg	cggcgcatcg	ctggagccgc	tgcttgccagc	4260
	atagcccttg	ggaagtgctc	tgtggatat	ttggcaggtt	atggagcagc	ggttgccagcg	4320
	gcgcctcgtg	ctcttaaggt	catgagccgc	gagttgccct	ccaccgagga	ctctggttaac	4380
	ctactccctg	ctatctctctc	ccctggcgcc	ctagtctcgt	gggttgctgt	cgcagcgata	4440
	ctcgctcgcg	acgtggagcc	aggggaaggg	gctgtcagtg	ggaagaacgc	gttgatagac	4500
	ttcgtttcgc	ggggtaacgc	cgtctccccc	agctactgat	tgctgtgagc	cgaactctga	4560
15	gcacagatca	ctcagatcct	ctctagtctt	accatcactc	agctgtctga	gagccttca	4620
	cagtggatca	acgagagctg	ctccacgcga	tgctcccgct	cgtggctaa	agatgttttg	4680
	gattggatat	gcacggtgtt	gaactgattc	aagacctgac	tcacgtccaa	gctcctgcgc	4740
	cgttgtccgg	gagtcctcctt	ctctctcatg	caacgtgggt	acaaagggat	ctggcggggc	4800
	gcagggcatca	tgcaaaaaca	ctgcacctgt	ggagcacatg	tcacccggca	tgtagaaaac	4860
20	ggttccatga	ggatcgtggg	gcctaggaac	tgtagttaac	cgtggatgag	aaactctccc	4920
	alttaacgctg	acacggcgct	ccctcccgcg	ccctcccgcg	cgcccaatga	gagcctctga	4980
	ctctggcggg	taggtctcga	ggagtcactg	gaagttacgc	gggttgggga	tttccaactc	5040
	tgacagggca	tgacacactg	caacgttaag	tgcccgctgc	aggttcocgc	ccccgaattc	5100
	ttcacagaag	tggtatgggt	gcggttgca	aggttaogctc	cagcgtgcaa	accctctcta	5160
	cgggaggaag	tcacattctc	ggtcgggctc	aatcaataac	tggttggtgc	acagctccca	5220
	tgagagcctg	aacccgagtg	agcagtgctc	acttccatgc	tcaccgaccc	ctcccaactt	5280
25	acggcgagga	cggctaaagg	taggtctggc	aggggactc	cccccccttt	ggccaagctc	5340
	tcagctagcg	agctgtctgc	gcttctctg	aaagcaaat	gcactaacgc	tcctgtacac	5400
	cggcgctcgt	acctcctcga	ggcccaacct	ctgtggcgcc	aggaagtggc	ggcgcaactc	5460
	acccgcgtgc	agtcagaaaa	taaggtagta	attttggaat	ctttcgagcc	gctccaaagc	5520
	gaggaagatg	agagggaggt	atccgtttcc	ggcgagatcc	tgccggaggtc	caggaatttc	5580
	ctccagagca	tgcccatatg	ggcaccgccg	gattacaacc	ctccactggt	agagctcgtg	5640
	aaggaccccg	actacgtccc	tccagtggta	caagggtgtc	cattgcgcgc	tgccaaggcc	5700
30	ctcccgatac	ccacttccag	gaggaagagg	acgggtgttc	tgctcgaact	tacgctgtct	5760
	ttgtcctctg	ggagctcgcc	cacaggaacc	ttccgcagct	cogaaatgct	ggccgctcag	5820
	acgggcaacg	caaagcgtct	tccgtgcacg	ccctcccgag	acgggaaagc	gggatccgac	5880
	gttgagctgt	actcctccat	gcctccctct	gaaggggagc	cggggagatcc	cgatctcagc	5940
	gaocggctct	ggctcaccgt	aagcagagag	gctagtgaag	aogtctctct	ctgctcgatg	6000
	tcctcacatc	ggacagggcg	ctcgtatcag	ccatgocgtg	cggaggaaac	caaagctccc	6060
	atcaatgcac	tgagcaactc	tttgctccgt	caacacaact	tggtctatgc	tacaacatct	6120
35	cgcagcgcaa	acctgcggca	gaagaaggtc	acctttgaca	gactgcaggt	cctggacgac	6180
	cactcacogg	acgtgctcaa	ggagatgaag	gcgaaggcct	caacagttaa	ggcaaaactt	6240
	ctatccctgg	aggaagctgt	taagctgagc	cccccaactt	cggcgagatc	taaatgtgtc	6300
	taatgggcaa	aggaagctcg	gaacctatcc	agcaaggccg	ttaacacact	ccgctcgctg	6360
	tggaaggact	tgctggaaag	cactgagaca	ccaattgca	ccaactcat	ggcaaaaat	6420
	gaggtttttc	gcgtccaaac	agagaagggg	ggccogaagc	cagctcgcct	tatcgtattc	6480
40	ccagattttg	gggtttcgtg	gtgogagaaa	atggcccttt	acagatgtgt	ctccaccttc	6540
	ctccagcgcc	tgatgggctc	ttcatacggg	ttccaatact	ctcctggaca	gcgggtcgag	6600
	ttcctgttga	atgctctgaa	agogaagaaa	tgccctatgg	gcttcgata	tgacaaccgc	6660
	tggttttgact	caacggtcac	tgagaaatgc	atccgttgtt	agggatcaat	ctaccaatgc	6720
	tgtagcttgg	ccccgaagc	cagcaacatc	ataagtgctg	ccagagagc	gctttacatc	6780
	gggggcccct	tgactaatc	taaggggcag	aactcggctg	atcccgctgt	cccgccagac	6840
	gggtgactga	cgaacagctg	cggtaataac	ctcacatggt	acttgaagcg	cgtctcgccc	6900
45	tgctcagctg	cgaagctcca	ggactgcacg	atgctcgtat	gcggagagaca	ccctgtcgtt	6960
	attcgtgaaa	gcggggggag	ccaagaggac	gagggagagc	tacgggcctt	cacggaggct	7020
	atgactagat	actctgcccc	ccctggggac	ccgcccaaac	cagaatacga	cttggagttg	7080
	ataacatcat	gctcctccaa	tgctgcagtc	gcgcagatg	catctggcaa	aagggtgtac	7140
	tattcacacc	gtgacccccc	acccccctgt	gcgcgggtgt	gttggagagc	agctagagac	7200
	actccagcca	attcctggct	agcaacatct	atcagtgatg	ccgcccactt	gtgggcaagg	7260
	atgactccta	tgactcatct	cttctccatc	cttctagctc	aggaacaact	tgaaaaagcc	7320
50	ctagatgtgc	agatcaacgg	ggcctgttgc	tcacatlgac	cacttgacct	acctcagatc	7380
	attccaagac	tcacatggct	fagcgcattt	taactccata	gttactctcc	aggttgagatc	7440
	aataggtgtg	cttcatcgct	caggaaactt	ggggtaacgc	ccctgcgagc	ctggagacat	7500

EP 1 043 399 A2

```

5      cgggccagaa gtgtccgcgc taggtactg tccaggggg ggagggtctgc cacttggtgc 7560
      aagtacctct tcaactgggc agtaaggacc aagctcaaac tcactccaat cccggctgcg 7620
      tcccagttgg atttatccag ctggttcggt gctggttaca gcgggggaga catatatcac 7680
      agcctgtctc gtgcccgacc ccgctggttc atgtggtgcc tactcctact ttctgtaggg 7740
      gtaggcattc atctactccc caaccgatga acggggagct aaacaactcca ggccaatagg 7800
      ocatcctgtt tttttccctt ttttttttc ttttttttt ttttttttt ttttttttt 7860
      ttttctcctt tttttttcct ctttttttcc ttttctttcc tttggtggct ccatcttagc 7920
      cctagtcaag gctagctgtg aaagggtcgt gagccgcttg actgcagaga gtgctgatac 7980
10     tggcctctct gcagatcaag t                                     8001

```

15

20

25

30

35

40

45

50

55

<210> 11
 <211> 11076
 <212> DNA
 <213> Hepatitis C Virus

<400> 11
 gccagccccc gattgggggc gacactccac catagatcac tcccctgtga ggaactactg 60
 tcttcacgca gaaagcgtct agccatggcg tagtatgag ttctgtgcag cctccaggac 120
 ccccctccc gggagagcga tggcagcttg cggaaacccg gactacaacg gaattgtcac 180
 gacgacoggg tctttctctg gatcaacccg ctcaactgct ggagatttgg gcgtgcccac 240
 g-cgagactgc tagccagta gtgttgggtc gcgaaagcg ttgtgtact cctcgatagg 300
 gtgcttgcca gtgcgccggc aggtctccta gcccgcgcac catgagcacg aatcctaacc 360
 ctcaagaata aaccacaagt aacaccaaag gcgcgcccac gattgaacaa gatggattgc 420
 acgcaggctc tccgcgcgct ttgggtggga ggctatttcg ctatgactgg gcacaacaga 480
 caatcgctcg ctctgatgcc gcgctgttcc ggctgtcagc gcaggggcgc cgggttcttt 540
 ttgtcaagac cgactcgtcc ggtgcccctga atgaactgca ggaacgagca gcgcgcctat 600
 cgtggctggc cagcagcggc gttccttgcg cagctgtgct gacgttcttc ctgaaagcgc 660
 gaagggactg gctgctattg ggcgaagctg cggggagga gctctgtcca tctcaacttg 720
 ctcttcgcca gaaatgatcc atcatggctc atgcaatgcg gcgcgtgcac acgcttgatc 780
 cggctacactg occattgcac caccagcgca aacatcgcat cgagcgagca cgtactcgga 840
 tggaaagccgg tcttctcgat caggatgata tggacaagaa gcaatcaggg ctgcgcgcag 900
 cgcaactgtt ccgcaagctc aagcgcgcga tgcccgaagg cgaggatctc gtctgacccc 960
 atggcgatgc ctgcttgccg aatatcatgg tggaaaaatg ccgcttttct gattcatcgc 1020
 actgttgccc gtctgggtgtg gcggaacgct atcaggacat acgcttgctt acccgtgata 1080
 ttgtcgaaag cgttgggcgc gaatggggct acgcttctct cgtgctttac ctgacacgpc 1140
 ctcccgatcc gacgcgcttc cctctctatc gcttcttcga cgtgttcttc tggattttaa 1200
 cagaccacaa cggtttccct ctagcgggat caattccgac cctctccctc cccccccctt 1260
 aaogttactg gccgaagccg cttggaaata ggccggtgtg cgttttgtata tatgttattt 1320
 tccaccatat tgcgctcttt tggcaatgtg agggcccgga aaacctggccc tgtcttcttg 1380
 acgagcattc cttaggggtct tccccctctc gccaaaggaa tgaacgattt gttgaatgtc 1440
 gtgaaggaa cagtctctct ggaagctctc tgaagacaa caactctctg agcagacctt 1500
 tgcagggcag ggaacccccc acctggcgac aggtgcctct gcggccaaaa gccacgtgta 1560
 taagatacac ctgcaaaagg ggcacaaacc cagtgccag ttgtgagtgt gatcagtgtg 1620
 gaagagtgca aatggctctc ctcaagcgta ttcaacaagg tctcgaagga tgcacgaag 1680
 ctacccattt gtatgggata tgatctgggg cctcoggtgca catgctttac atgtgtttag 1740
 tcgaggttaa aaaaactgta gccccccgca accacgggga cgtgtgtttc ctttgaaaaa 1800
 cagcataata ccatggggcac gaatcctaaa cctcaaaaga aaaccaaacg taacaccaac 1860
 cgcgcgccac aggaactcaa gttcccgggc ggtggtcaga tctcgtgtgg agtttacctg 1920
 ttgcgcgcga ggggcgccag gttgggtgtg cgcgcgacta ggaagacttc cgagcggtcg 1980
 caactcgttg gaagcgacaa acctatcccc aagctctccc agcccgaggg tagggcctgg 2040
 ctgcagcccg ggtacccctg gccctctcat ggcaalgagg gcttggggtg ggcaggaatg 2100
 ctccgttcac ccgtggtctc tccgctcatt tggggcccaa cgacaccccg cgttagtgct 2160
 cgcaatttgg gtaagctcat cgataccctc acgtgcgctc tccgcgactc catggggatc 2220
 attcgcctcg tggcgccccc cctagggggc gctgccaggc ccttgcgcca tggcgtccgg 2280
 gttcttgagg acggtcgtaa ctatgcaaca gggaaatctc ccggttgctc cttttctatc 2340
 ttctttttgg ctttgcgttc ctgtttgacc atccagactt ccgcttatga agtgcgcaac 2400
 gtatccggag tgcattcgt cagcaacgac tgcctcaacg caagcattgt gtatgagcca 2460
 gggcacaatga tcatgcatac ccccggtgct gtgcctcgcg ttcggagaaa caactctccc 2520
 cgtgtctggg tagcgtctac tcccacgctc gggcccgaga acgctagctt ccccacagt 2580
 acgatacag gccatgtgca ttgctcgttt gggcggtgtc cttctgtctc cgttatgta 2640
 gtggagatct tctggagatc tgttttccct gtgcgcagag ttgttaacct ctgcgctcgc 2700
 cggcacgaga cagtacagga ctgcaattgc tcaatatc cccgcgcact gacaggtcac 2760
 cgtatggctt gggatatgat gatgaactgg tcaactacag cagccctagt ggtatcgacg 2820
 ttaactccga tcccaaacg tctcgtggat atggtggcgg gggccattg gggagtctca 2880
 gcggcgcttg cctactatct catggtgggg aactgggcta aggttctgat tgtgatgta 2940
 ctctttgcgg cgttgtaacg gggaaacctat gtgacagggg ggaagatgac caaaacacc 3000
 ctccggatta cgtccctctt ttaccccggg tcatcccgga aaatcagctt tgtaaacccc 3060
 aacgacagct ggcatatcga caggactcgc ctgaactcct atgactccct caaacctcgg 3120
 ttctttgtgt cgtgtctcta cgtgcacaa ttcaactcct ctgagtgccc agagcgatga 3180
 gccagctgca gcccatatga cgtgttgcct caagggtggg ggcccatcac ttacatagg 3240
 tcaacagct cggaccagag gccttattgt tggcactac caccocggcc gtgcggtatc 3300
 gtaccccggg cgcaggtgtg tgttccagtg taactgttca ccccaagccc tctcgttggt 3360
 ggaacgacgg accggttccg cgtccctacg tacaagtggg gggagaaTga gacggacgtg 3420
 ctgctctcta caaacacgcg gccgcgcgca ggaacatggt ttgctgtcat atggtatgat 3480
 agcatggggt tccacaagac gtgcgggggc ccccgtgta acatcggggg gatcgccaat 3540

5
 10
 15
 20
 25
 30
 35
 40
 45
 50
 55

aaaaccttga cctgcccac ggaactgttc cggaaagacc cggaggccac ttacaccaga 3600
 ttggtgttcgg ggccttgggt gacacacaga ttgcttgctc ataccacata caggtcttgg 3660
 cactacccctt gcactgtcca cttaccactc ttcaaggtta ggaatgaact ggggggagtg 3720
 gagcaacaggc tcgaagccgc atgcaattgg actcgaggag agccttgtaa cctggaggac 3780
 agggagacagt cagagcttag ccogctgtcg ctgtctacaa cggagtggca ggtattgcc 3840
 tggctcttcca ccaacctacc ggcctgtccc actggtttga tccatccca tcagaagctc 3900
 gtggactgtac aatacctgta cgtataaggg tcggcggttg tccctcttgc aatcaaatgg 3960
 gggatgtctcc ttgtgtctctt cctctctctg cggagcgctg gctctgtgct ctgcttgggg 4020
 atgattgtcgc tgaatgtcca agctgagaga acctgtgtgt cctcaacgtg cctcaacgtg 4080
 gcalcctgag cggggggcca tgcattcttc tccctctctg tggctctctg tgcctcctgg 4140
 taacatcaagg cgaagctggc cccgtggggc gcatatgccc tctacgggt atggccgcta 4200
 ctccctgtccc ttgctggcgtt accacacaga gcatatgcca tggacccggga gatgcagca 4260
 tctgtcggag gcgcggtttt cgtaggtctg atactcttga atactcttca tcaaccgggc cgaggccac 4320
 ctgctccctcg ctaggctcat atggttggtta caattattta tcaaccgggc cgaggccac 4380
 ttgcaagtgt ttgtcccccc cctcaacgtt cggggggggc ggaatgcgt catctctctc 4440
 acglgcggca tccaccacaga gcatattctt accatccca aaactctgtc cgcatactc 4500
 ggtccactca ttgtgtccca ggtcgttata accaagtgcc cgtactctgt cgcgcacac 4560
 gggctcattc gtgcattgat gctggtgcgg aagttgtctg ggggtcatta tgtccaagt 4620
 gctctcatga agttggccgc actgacaggt acgtactgtt atgacatct caccocactg 4680
 cgggactggg cccacggggc cctacagagc ctgtcggttg cagttgagcc cgtcgtctc 4740
 tctgatatgg agaccaaggt tatcaactgg ggggcagaca ccgcgcgtg tggggacac 4800
 atcttggggc ttgcctgtct cgcgcgagg ggaagggaga tacatctgg acgcgcagac 4860
 agccttgaag ggcagggggt gcaactctcc gcgcattata cggctactc ccaacagagc 4920
 cagagctctac ttgcgtctct caaccagggc ctcaacgggc gggacagaga cgcgactc 4980
 gggagggagcc aagtgtctc caaccacaca caactcttcc tggcgactgt cgtcaatgac 5040
 gtgtgttga ctgtctatca ttgtgcggcc tcaaaagccc ttgcgcggcc aaaggggcca 5100
 atcacccaaa tgtacaccaa ttgtgacagc gaactcgtcg gctggcaagc gcccccggg 5160
 gcgcgttctc tgacaccatg caactcgccc agctcggacc tttactgtgt cagcagcatc 5220
 gcgatgtgca ttccggtgag cggcgggggc gacagcagg ggaactact ctccccagg 5280
 cccgtctctc acttgaagg ctctctgggc ggtccactgc tctgcacctc ggggcagct 5340
 ttgggactct ttggcgctg cgttgcacc cgggggttg cgaagcgtt ggaacttga 5400
 cccgtcaggt ctatgggaac catgacggg ttcccggttc tccagggaca ctgcctact 5460
 cggcgctac cgcagacatt ccaagtgccc catctacgc cccctactg tagcggcaag 5520
 agcactaagg ttgcgcgtcg gtatgcaggc caagggtata agtgctgtg cctgaacccc 5580
 tccgtgcgcc ccaccttagg ttccggggcg tatatgtcta aggcacatg tatcgaacct 5640
 aacatcagaa tcggggtaag gaocatcacc acgggtggcc ccatcagta ctccacctat 5700
 ggcgaagtctc ttgcgcagg ttggtgtctct ggggcgcctc atgacatcat aatatgtat 5760
 gagggtccat caactgactc gacacatcc ttggcgatcg ggcagctcat ggaacaggc 5820
 gagcggctg gagcggact cgtcgtctc gccacgctc cgcctcggg atcgtacc 5880
 gtgccacatc caaatacga ggaagtggt ctgtccagca ctggagaat cccctttat 5940
 ggcaaaagcca tccocatga gacatcaag gggggaggc acctcattt ctgcacttc 6000
 aagaagaat gtgatgagc cgcgcgaag ctgtccggcc tcggactcaa tgcgttagca 6060
 tattacoggg gccttgatgt atcgtcata ccaactagcg gacgctcat ttgcttagca 6120
 aoggacgctc taatgacggg cttaccggc gacttogact catgatoga ctgcaatca 6180
 ttgttcaacc agacagtga cttcagcctg gaccacact tcaacttga gacgacacc 6240
 gtgcacaag acgggtgtc acgtgcag cggcagga gactgtgtg gggcagagt 6300
 ggcatttaca ggttgtga tccagagaa cggctctgg gcatgtga tctctcgtt 6360
 ctgtgcgagt gcatgacgc ggcctgtgct ttgttcagc tcaacggc cgagacctca 6420
 gttaggttgc ggccttacct aaacacacca gggttgccg tctccagga ccatctggag 6480
 ttctggggaga gctctttac aggcctcacc cacatagag ccactttct tcccaagct 6540
 aagcaggcag gagaacatt cccctacctg gtatgatacc aggtcaoyt gtgcgccag 6600
 gctcaggctc caactccatc gtgggaccaa atgtggaa gtctctacg gctaaagct 6660
 agctgtcagc ggcacaagcc cctgctgat agctgggag ccgttcaaa cagagttact 6720
 atcacacacc cctacacaa atactatcc gcatcattc cgtctgact cggagctgt 6780
 acgagcactc ggtgtcgtgt aggcgaagc ctacagactc tggccggtt ttgcttga 6840
 acaggcagc ttgtcattgt ggcggagatc atcttctcc gaaagccgc catctctcc 6900
 gcaggggaag tccattaccg ggaattcga gataggaa agtcgcctc acactctcc 6960
 tacatcgaa aggaatgaa gctgcgcgaa caattcaac agaagcaat cgggttctg 7020
 caaacagcca ccaagcaagc ggaagctgct gctccgtg ttgaaatcaa gtggcgacc 7080
 atcgaaacct tctggggcaa gcatatgtgt aattcaatc tggggatata atatttga 7140
 ggttgttcca cctgctctgg caaccggcg atagcatcac tgatgggatt cacagctct 7200
 atcacagcg cctcacac caaagtatc cctctgttga acatctggg ggaatgggt 7260
 gcgcaccaac ttgctctccc cagcgtcgt ttgtgttgg taggcggcg catcgtga 7320
 gcgctgttg gcagcatagg ccttgggaag gtgctgtgg atattttgc aggttatga 7380
 gcagggggt caggcgctc cgtggcctt aaagtcatga cggcgagtt gccctccac 7440
 gaggacctg ttaactact cctgtatcc ctctccctt ggccttagt cgtcgggct 7500

	gtgtgctggcag	cgatactcg	tcggcactgt	ggccacgggg	agggggcgtgt	gcagtggatg	7560
	aacgggtgta	tagcgttccg	ttcgccgggt	aaccacgtct	ccccacgga	ctatgtgctt	7620
	gcagacgacg	ctgcagtcac	agtcaactcg	atccctctca	gtcttaacct	caactagcgt	7680
5	ctgaaggcgc	ttccaccagt	gatacaacga	gactgtctcc	cgccatgtct	cggtgtgtgg	7740
	ctaagagatg	tttgggagt	gatatgcacg	gtgtgactct	atttcaacag	ctgtgtcctg	7800
	tcacaagctcc	tgcccgcatt	gccggagatc	ccctcttctc	catgtcaacg	ctggctacaag	7860
	ggagtgctggc	ggggcgacgg	catcatgcac	acccctcgcc	catgtggagc	acagatcac	7920
	ggacatgtga	aaaacggttc	catgagatc	gtggggcggt	ggacctgtag	taacacggtg	7980
	catggaaact	tcctccattaa	cgctacaccc	acgggcccct	gcacgcccct	cccgggcgca	8040
10	aattatttota	gggcgcgtgt	gcgggttagt	gctgagagat	acgtgagggt	tacggcgctg	8100
	ggggattttcc	actacgtgac	gggcatgacc	actgacaa	ttaaagtccc	gtgtcaggtt	8160
	ccgggcccgcg	aattctctac	agaagtggat	ggggctcggt	gtccacaggt	cgctccacgc	8220
	tgcaaacccc	tcctacggga	ggaggttcaca	ttctctgctg	ggctcaatca	atacctggtt	8280
	gggtcacagc	tcctatgcga	gcctgaacgc	gatgtacag	tgctcacttc	catgtccacc	8340
	gacccctccc	acattacggc	ggagacggct	aaagctaggt	tgggcagggg	atctcccccc	8400
15	cccttgggcca	gctcatcagc	tagccagctg	ttctgcctct	ccctgaagc	aacatgcact	8460
	accctgcatg	actccccgga	gcctgacctc	atcgaggcca	acctctgtgt	ggcgacaggg	8520
	atggggcgga	acatcaacgc	cgtagagtc	gaataaagg	ttagtaattt	ggactctctt	8580
	gagccgctcc	aagcggagga	ggatgagagg	ggagatccg	tgccggcgga	gatcctggcg	8640
	aggtccaggga	aattccctcg	agcgatgccc	atatggcgac	gcccgagta	caaccctcca	8700
	ctgttagagtt	cctggaagaa	ccccgactac	gtccctccag	tggtacacgt	gtgtccattgt	8760
20	ccgecttgcca	agggcccctcc	gatcaaacct	tcacggagga	agaggacgtg	gtctctgtca	8820
	gactctaacg	tgctctctgc	ctggggagag	ctcgccagca	agactcttcg	agctccgaa	8880
	tcgtcgccgc	tcgaacaggg	cacggcaacg	gcctctctgt	accagccggc	cgagccgggc	8940
	gaacggggat	ccgacgttga	gtcgtactcc	tcactgcccc	cccttgaggg	ggagcgccgg	9000
	gtaccggatc	tcagcgacgg	gtcttggtct	accgtgaagc	agggagctag	tgaggacgtg	9060
	gtctgtctgt	cgatgtccta	cacatggaca	ggcgccctga	tcacgcctag	cgctgcggag	9120
25	gaacaacagc	tgccccatca	tgcactgagc	aactctttgc	tcctgcacca	caactgtgtc	9180
	tatgttaca	catctcgacg	cgcaaacctg	cgccagaaga	aggtcaacct	tgacagactg	9240
	caggtctctg	acgaccacta	ccgggacgtg	ctcaaggaga	tgaaggcgaa	ggcgctccca	9300
	gttaaaggcta	aactctctac	cgtggagga	gcctgtaagc	tgaagccccc	acattgccgc	9360
	agatctaaat	ttggctatgg	ggcaaaaggac	gtccggaaac	tatccagcaa	ggccgttaac	9420
	cacatccgct	ccgttgaggaa	ggaactgtgt	gaagacactg	agacaccaat	tgacacacac	9480
30	atcatggcaa	aaaatgaggt	tttctcgctc	caaccagaga	agggggggcg	caagccagct	9540
	cgccctatcg	tattccccga	tttgggggtt	cgtgtgtcgc	agaaaatggc	ccctttacgt	9600
	gtggtctcca	ccctccctca	ggccgtgatg	ggctctctac	acggattcca	atactctcct	9660
	ggagacgggc	tcaggttctc	ggtgaatgcc	tgaaaacga	agaaa tgccc	tatgggcttc	9720
	gcataatgaca	cccgtgtttt	tgactcaacg	gtcaactgaga	atgacatccg	tggttgagag	9780
	tcaactcacc	aatggttgga	cttggccccc	gaagccagac	agggccataag	gtcgtccaga	9840
35	gagcgcgttt	acatcggggg	ccccctgact	aattctaaag	ggcgaaactg	cggtctatcg	9900
	cggtgccgcg	cgagcgggtg	actgaacgac	agctgcggta	ataccctcac	atgttaactgt	9960
	aaaggccgtg	cgccctgtcg	agctcggaag	ctccaggagc	gcacgatgct	cgatatcgga	10020
	gacgaccttg	tcgttatctg	tgaaggagct	ggagcccaag	agagcagggc	gagcctacgg	10080
	gccttcacgg	agcctatgac	tgaatactct	gcccccctcg	gggacccggc	caaacacgga	10140
	tacgacttgg	agttgataac	atcatgtccc	tccaatgtgt	cagtccggca	cgatgatcct	10200
40	ggcacaaggg	tgtaactatc	acccgtgac	ccccaccccc	cccttgccgcg	ggctgctgtg	10260
	gagacagcta	gacacactcc	agtcacttcc	tggtcaggca	acatcatcat	gtatgcggcc	10320
	accttgtggg	caaggatgat	cctgatgact	catctttctc	ccatcctctc	agctcagaa	10380
	caacttgaaa	aagccctaga	ttgtcagatc	tacggggcgt	gttaactcat	tgagccactt	10440
	gacctacctc	agatcaattca	acgactccat	ggccttagcg	ctttttcact	ccatagttaa	10500
	ctctccaggtg	agatcaaatg	ggtgggttca	tgccctcagga	aacttggggt	acggccctgt	10560
45	cgagtctgga	gacatcgggc	cagaagtgtc	cgcgctagac	tactgtccca	ggggggggag	10620
	gctgcccactt	gtggcaagta	ccctctcaac	tgggcagtaa	ggccaaagct	caaacactca	10680
	ccaatcccg	ctcgttccca	gttgagatta	tcaagctgtg	tcagctcgtg	ttacagcggt	10740
	ggagacatat	atacagcctc	gtctcgtgcc	cgccccacac	ggttctatgt	gtgcctcact	10800
	ctactttctg	taggggttag	catcatatca	cccttttttt	gatgaacggg	gagctaaaca	10860
	ctccagggcca	ataggccatc	ctgttttttt	cccttttttt	ttttcttttt	tttttttttt	10920
50	tttttttttt	tttttttttc	tccttttttt	ttctcttttt	tttctctttc	tttctctttc	10980
	tggtctccatc	ttagccctag	tcacgctcag	ctgtgaaagg	tcctgtgagcc	gcttgactgc	11040
	agagagtgct	gatactggcc	tctctgcaga	tcaagt			11076

Patentansprüche

1. Hepatitis C Virus (HCV) Zellkultursystem, das im wesentlichen eukaryontische Zellen umfaßt, die eingeschleustes HCV-spezifisches Genmaterial enthalten, dadurch gekennzeichnet,

5

daß die eukaryontischen Zellen humane Hepatomazellen sind und daß das eingeschleuste HCV-spezifische Genmaterial ein HCV-RNA-Konstrukt ist, das die HCV-spezifischen RNA-Abschnitte 5' NTR, NS3, NS4A, NS4B, NS5A, NS5B und 3' NTR und zusätzlich ein selektierbares Markergen (Selektionsgen) umfaßt.

10

2. Zellkultursystem nach Anspruch 1, dadurch gekennzeichnet,

daß die Hepatomazellen von einer handelsüblichen Hepatomazelllinie abstammen.

3. Zellkultursystem nach Anspruch 1, dadurch gekennzeichnet,

15

daß die Hepatomazellen aus einer Hepatoma-primärzellkultur gewonnen sind.

4. HCV-RNA-Konstrukt, dadurch gekennzeichnet,

20

daß es die HCV-spezifischen RNA-Abschnitte 5' NTR, NS3, NS4A, NS4B, NS5A, NS5B und 3' NTR und zusätzlich ein selektierbares Markergen (Selektionsgen) umfaßt.

5. HCV-RNA-Konstrukt nach Anspruch 4, dadurch gekennzeichnet,

25

daß es eine Nukleotidsequenz gemäß einem der Sequenzprotokolle SEQ ID NO: 1 bis SEQ ID NO: 11 umfaßt.

6. HCV-RNA-Konstrukt nach Anspruch 4, dadurch gekennzeichnet,

30

daß die 3' NTR eine Nukleotidsequenz aufweist, die aus der Gruppe der nachfolgend aufgelisteten Nukleotidsequenzen (a) bis (i) ausgewählt ist:

(a) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTTTTTT
TTAGCTTTTTTTTTTCTTTTTTTTGAGAGAGAGAGTCTCACTCTG
TTGCCAGACTGGAGT

35

40

45

50

55

- 5 (b) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTT
TTTTTAGTCT TTTTTTTTTC TTTTTTTTGA GAGAGAGAGT
CTCACTCTGT TGCCAGACT GGAGC
- 10 (c) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTT
TTTAATCTTT TTTTTTTTCT TTTTTTTTGA GAGAGAGAGT
CTCACTCTGT TGCCAGACT GCAGC
- 15 (d) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTT
TTTTTAGTCT TTTTTTTTTT TCITTTTTTT TGAGAGAGAG
AGTCTCACTC GTTGCCAG ACTGGAGT
- 20 (e) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTT
TTTTTAGTCT TTTTTTTTTT TCITTTTTTT TGAGAGAGAG
AGTCTCACTC GTTGCCAG ACTGGAGT
- 25 (f) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTT
TTTTTAGTCT TTTTTTTTTT TCITTTTTTT TGAGAGAGA
GAGTCTCACT CTGTTGCCA GACTGGAGT
- 30 (g) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTT
TTTTTAGTCT TTTTTTTTTT CTITTTTTTT GAGAGAGAGA
GTCTCACTCT GTTGCCAGA CTGGAGT
- 35 (h) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTT
TTTTTTAAT CTITTTTTTT TTTTCCTTT TTTTGAGAGA
GAGAGTCTCA CTCTGTTGCC CAGACTGGAG T
- 40 (i) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTT
TTTTTAATC TTTTTTTTTT TTICTTTTT TTTTGAGAG
AGAGAGTCTC ACTCTGTTGC CCAGACTGGA GT

- 45 7. HCV-RNA-Konstrukt nach einem der Ansprüche 4 bis 6, dadurch gekennzeichnet,
daß das selektierbare Markergen ein Resistenzgen und insbesondere eine Antibiotikaresistenzgen ist.
- 50 8. HCV-RNA-Konstrukt nach einem der Ansprüche 4 bis 6, dadurch gekennzeichnet,
daß das selektierbare Markergen ein Neomycinphosphotransferasegen ist.
- 55 9. HCV-RNA-Konstrukt nach einem der Ansprüche 4 bis 8, dadurch gekennzeichnet,
daß das selektierbare Markergen strangabwärts der 5' NTR in die HCV-RNA integriert ist.
10. HCV-RNA-Konstrukt nach einem der Ansprüche 4 bis 9, dadurch gekennzeichnet,

daß das selektierbare Markergen über ein Ribozym bzw. eine Erkennungsstelle für ein Ribozym mit der HCV-RNA verbunden ist.

11. HCV-RNA-Konstrukt nach einem der Ansprüche 4 bis 10, dadurch gekennzeichnet,

daß es ein integriertes Reportergen aufweist.

12. HCV-RNA-Konstrukt nach Anspruch 11, dadurch gekennzeichnet,

daß das Reportergen ein Gen aus der Gruppe der Luziferasogene, dem CAT-Gen (Chloramphenicol-Acetyltransferase-Gen), dem lacZ-Gen (beta-Galaktosidasegen), der GFP-Gene (green-fluorescence-protein-Gene), dem GUS-Gen (Glukuronidasegen) und dem SEAP-Gen (Sekernierte-Alkalische-Phosphatase-Gen) ist.

13. HCV-RNA-Konstrukt nach einem der Ansprüche 4 bis 11, dadurch gekennzeichnet,

daß deren Replikation die Expression eines (zellulären) Surrogatmarkergens beeinflußt.

14. HCV-RNA-Konstrukt nach einem der Ansprüche 11 bis 13, dadurch gekennzeichnet,

daß das Resistenzgen derart in das offene Leseraster der HCV-RNA einkodiert ist, daß es erst nach einer proteolytischen Prozessierung in eine aktive Form überführbar ist.

15. HCV-RNA-Konstrukt nach einem der Ansprüche 11 bis 14, dadurch gekennzeichnet,

daß das Reportergen und das selektierbare Markergen derart räumlich in dem Konstrukt angeordnet sind, daß sie gemeinsam ein Fusionsprotein exprimieren.

16. Zellsystem nach einem der Ansprüche 1 bis 3 dadurch gekennzeichnet,

daß das HCV-RNA-Konstrukt ein Konstrukt gemäß wenigstens einem der Ansprüche 4 bis 15 ist.

17. Zellsystem nach Anspruch 1, dadurch gekennzeichnet,

daß das HCV-RNA-Konstrukt enthaltenden Zellen bei der DSMZ, Braunschweig, BRD, unter der Hinterlegungsnummer DSM ACC2394 (Laborbezeichnung HuBI 9-13) hinterlegt sind.

18. Verwendung eines Zellsystems nach einem der Ansprüche 1 bis 3 oder 16 bis 17 und/oder eines HCV-RNA-Konstrukts nach einem der Ansprüche 4 bis 15 zur Herstellung und/oder Evaluierung und/oder Testung von Therapeutika und/oder Diagnostika zur Behandlung von insbesondere HCV-Infektionen.

19. Verwendung eines Zellsystems nach einem der Ansprüche 1 bis 3 oder 16 bis 17 und/oder eines HCV-RNA-Konstrukts nach einem der Ansprüche 4 bis 15 zur Herstellung eines Impfstoffes gegen HCV-Infektionen.

20. Verwendung eines HCV-RNA-Konstrukts nach einem der Ansprüche 4 bis 15 zur Herstellung einer leberzell-spezifischen Gentherapie.

21. HCV-RNA-Konstrukt nach einem der Ansprüche 4 bis 15, dadurch gekennzeichnet,

daß es ein integriertes Fremdgen aufweist und dazu geeignet ist, dieses Fremdgen in eine Zielzelle einzuschleusen, die zur Expression dieses Fremdgens geeignet ist.

22. Verfahren zur Gewinnung von zellkultur-adaptierten Mutanten eines HCV-RNA-Konstrukts gemäß einem der Ansprüche 4 bis 15, wobei die Mutanten gegenüber dem HCV-RNA-Konstrukt eine erhöhte Replikationseffizienz aufweisen, dadurch gekennzeichnet,

daß man ein Zellsystem gemäß Anspruch 1, bei dem das eingeschleuste HCV-spezifische Genmaterial ein HCV-RNA-Konstrukt mit Selektionsgen nach einem der Ansprüche 4 bis 15 ist, auf/n dem dem Selektionsgen entsprechenden Selektionsmedium kultiviert, daß man die gewachsenen Zellsysteme erntet, und daß man

aus diesen Zellklonen die HCV-RNA-Konstrukte oder Teile davon isoliert.

23. Verfahren nach Anspruch 22, dadurch gekennzeichnet,

5 daß man die isolierten HCV-RNA-Konstrukte wenigstens einmal erneut passagiert, nämlich in Zellen eines Zellkultursystems nach Anspruch 1 einschleust, das dabei erhaltene Zellkultursystem gemäß Anspruch 1, bei dem das eingeschleuste HCV-spezifische Genmaterial das isolierte HCV-RNA-Konstrukt mit Selektionsgen ist, auf/in dem dem Selektionsgen entsprechenden Selektionsmedium kultiviert, die gewachsenen Zellklone erntet und aus diesen Zellklonen die HCV-RNA-Konstrukte isoliert.

10 24. Verfahren zur Herstellung von Mutanten eines HCV-Volllängengenoms oder eines HCV-Teilgenoms oder eines beliebigen HCV- Konstrukts mit im Vergleich zu dem ursprünglichen HCV-Volllängengenom oder -Teilgenom oder HCV-RNA-Konstrukt erhöhter Replikationseffizienz, dadurch gekennzeichnet,

15 daß man mit einem Verfahren nach Anspruch 22 oder 23 eine zellkultur-adaptierte Mutante eines HCV-RNA-Konstrukts herstellt und isoliert, daß man die Nukleotid- und Aminosäuresequenz dieser Mutante bestimmt und durch Vergleich mit der Nukleotid- und Aminosäuresequenz des ursprünglichen HCV-RNA-Konstrukts die Art, Anzahl und Positionen der Nukleotid- und Aminosäuremutationen bestimmt, und daß man diese Mutationen entweder durch gezielte Mutagenese oder durch Austausch von Sequenzabschnitten, die die betreffenden Mutationen enthalten, in ein (isoliertes) HCV- Volllängengenom oder ein HCV- Teilgenom oder ein beliebiges HCV-RNA-Konstrukt einführt.

25 25. Zellkultur-adaptiertes HCV-RNA-Konstrukt mit hoher Replikationseffizienz, dadurch gekennzeichnet,

25 daß es durch Nukleotid- und/oder Aminosäure -Mutationen von einem HCV-RNA-Konstrukt nach einem der Ansprüche 4 bis 15 ableitbar ist und daß es mit einem Verfahren nach einem der Ansprüche 22 bis 24 erhältlich ist.

30 26. Zellkultur-adaptiertes HCV-RNA-Konstrukt nach Anspruch 25, dadurch gekennzeichnet,

daß es einen oder mehrere der nachfolgend aufgeführten Aminosäureaustausche aufweist, nämlich 1283 arg -> gly und/oder 1383 glu -> ala und/oder 1577 lys -> arg und/oder 1609 lys -> glu und/oder 1936 pro -> ser und/oder 2163 glu -> gly und/oder 2330 lys -> glu und/oder 2442 ile -> val aufweist.

35 27. Zellkultur-adaptiertes HCV-RNA-Konstrukt nach Anspruch 25 oder 26, dadurch gekennzeichnet,

daß es einen oder mehrere der in Tabelle 3 aufgeführten Nukleotid- und/oder Aminosäureaustausche aufweist, wobei Tabelle 3 Bestandteil dieses Anspruchs ist.

40 28. Zellkultur-adaptierte Mutanten eines HCV-RNA-Konstrukts oder eines HCV-Volllängengenoms mit im Vergleich zu dem ursprünglichen HCV-RNA-Konstrukt oder dem ursprünglichen HCV-Volllängengenom erhöhter Replikationseffizienz, dadurch gekennzeichnet,

45 daß sie mit einem Verfahren erhältlich ist, bei dem man in einem zellkultur-adaptierten HCV-RNA-Konstrukt nach Anspruch 24 durch Sequenzanalyse und Sequenzvergleich die Art und Anzahl der Mutationen bestimmt und diese Mutationen in ein HCV-RNA-Konstrukt, insbesondere in ein HCV-RNA-Konstrukt gemäß einem der Ansprüche 4 bis 15, oder in ein (isoliertes) HCV-RNA-Volllängengenom einführt, entweder durch gezielte Mutagenese oder durch Austausch von Sequenzabschnitten, die die betreffenden Mutationen enthalten.

50 29. Hepatitis C Viruspartikel oder virus-ähnliche Partikel dadurch gekennzeichnet,

daß sie mit einem Verfahren nach einem der Ansprüche 22-24 erhältlich sind.

55 30. Zellen, infiziert mit Hepatitis C Viruspartikeln oder virus-ähnliche Partikeln gemäß Anspruch 29.

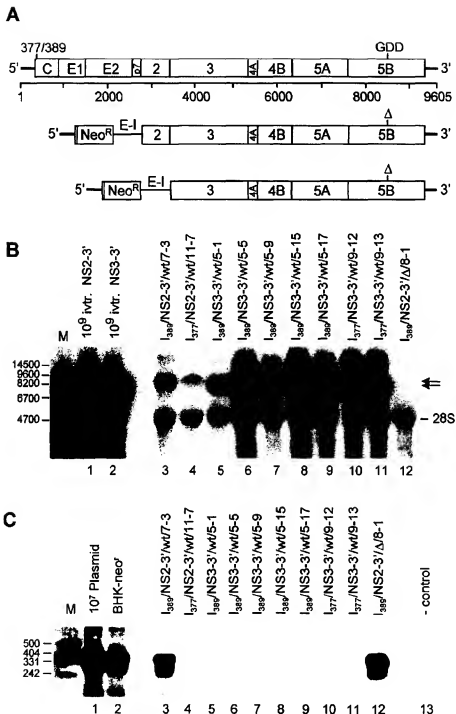
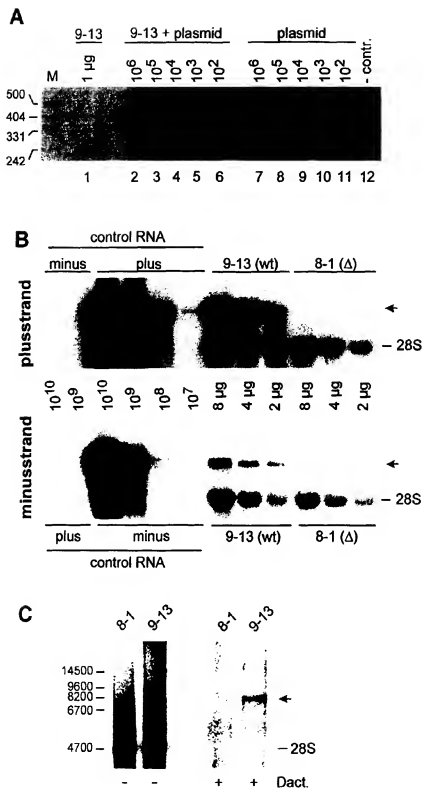


Fig. 1



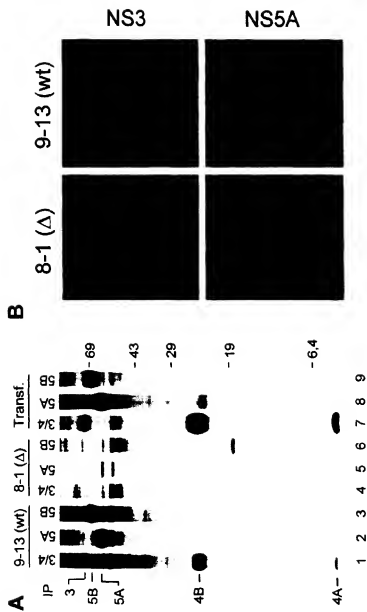


Fig. 3

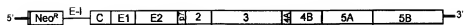
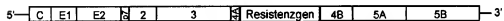
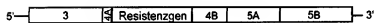


Fig. 4

A



B

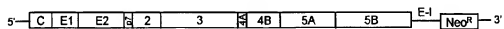
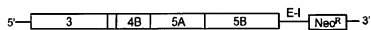


Fig. 5

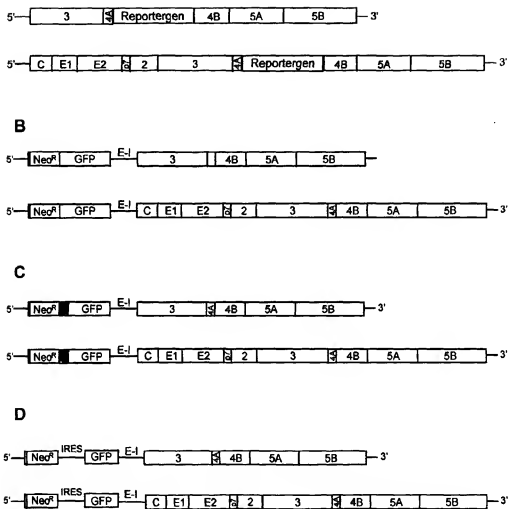


Fig. 6

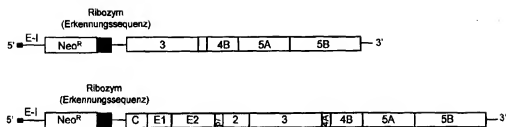


Fig. 7

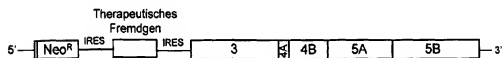


Fig. 8

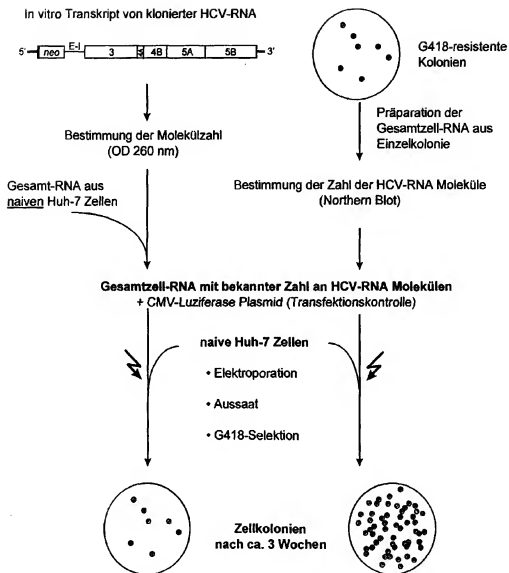


Fig. 9

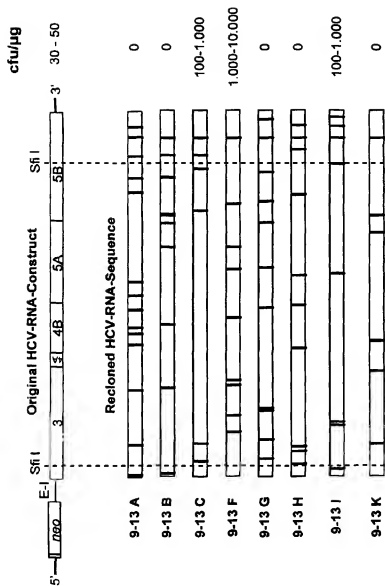


Fig. 10

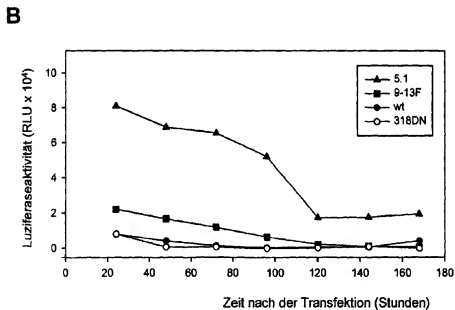
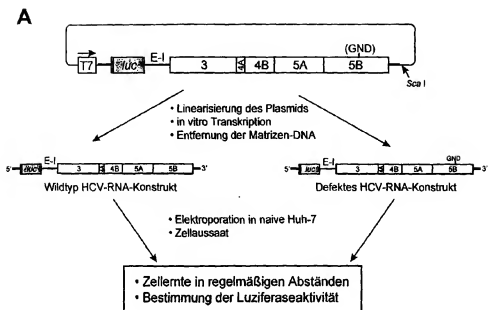


Fig. 11

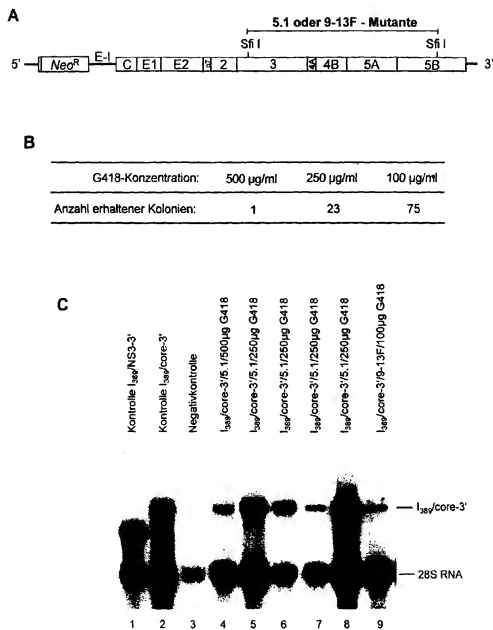


Fig. 12

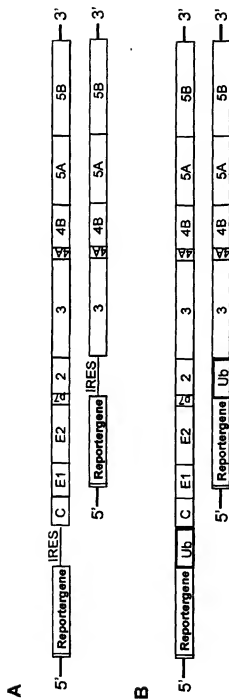


Fig. 13



Fig. 14



Fig. 15